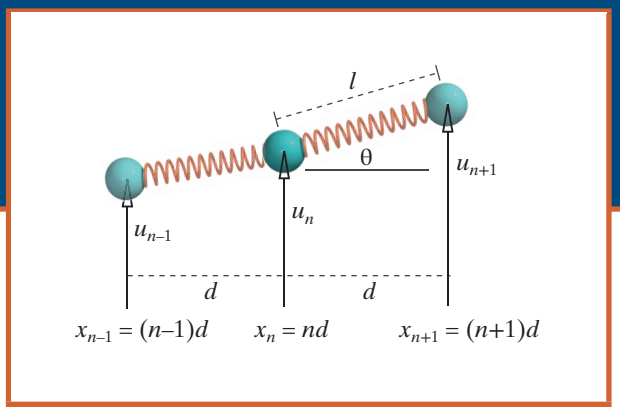
SUPPLEMENT

S3

Waves



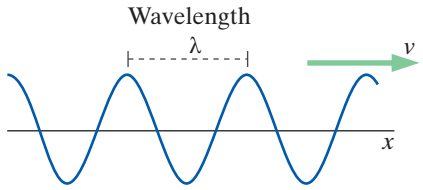


Figure S3.1 A sinusoidal wave propagating at speed v; the wavelength λ is the distance between crests of the wave.

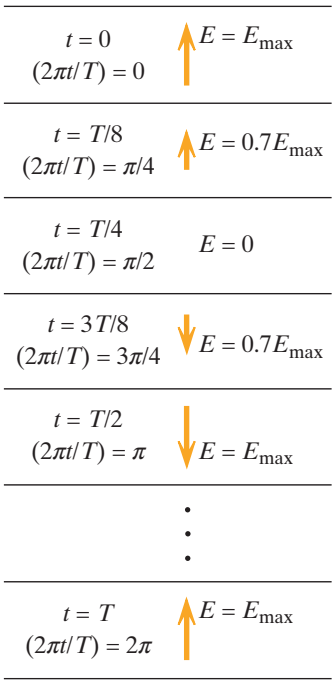


Figure S3.2 Snapshots of the electric field at one particular location, observed at successive times.

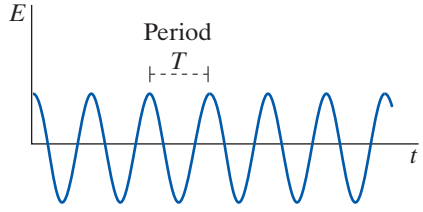


Figure S3.3 Graph of E_y vs. time t as observed at one particular location.

OBJECTIVES

After studying this supplement, you should be able to

- Determine the locations of maxima and minima in patterns of interference from two or more sources of light.
- Calculate the properties of longitudinal and transverse mechanical waves.
- Determine the patterns of "standing waves."
- Make connections between the particle model and the wave model of light.

S3.1 WAVE PHENOMENA

In a wave energy propagates from one place to another without the transfer of matter. Shake the end of a taut string and a pulse runs along the string, but the atoms in the string only move sideways (a transverse wave), not from one end of the string to the other. Hit a bar of aluminum at one end, and a pulse of sound propagates to the other end of the bar without any aluminum atoms moving from one end of the bar to the other (a longitudinal wave). Clap your hands, and your friend's ear is affected by the sound that propagated through the air, but your hands and the neighboring air stay with you (sound propagation in air is also a longitudinal wave). An electromagnetic wave can transfer energy and momentum to a charged particle. The refraction of light when it passes from one medium to another can be explained by considering light as an electromagnetic wave.

In the first part of this supplement we will examine additional phenomena that can be explained with a wave model of light, which include interference and diffraction of electromagnetic radiation, and reflection and scattering of light. Next we will investigate mechanical waves, such as waves on a string. Finally we will discuss the relationship between the wave model of light and the particle (photon) model discussed in Volume 1.

Variation in Time

If you stand at one particular location as a sinusoidal or cosinusoidal electromagnetic wave goes by (Figure S3.1), you observe the electric field varying in time like $E\cos(2\pi t/T)$. The time-varying quantity $(2\pi t/T)$ is called the *phase* of the cosine function. The phase $(2\pi t/T)$ increases by 2π rad (360°) whenever t increases by an amount T. Figure S3.2 shows as a series of snapshots what you would observe at various times, at one location in space, while Figure S3.3 shows the same observations as a graph of E_y vs. t.

Of course there is also a cosinusoidally varying magnetic field, at right angles to the electric field. We are concentrating on the electric field in electromagnetic radiation because it has a much bigger effect on matter than does the magnetic field.

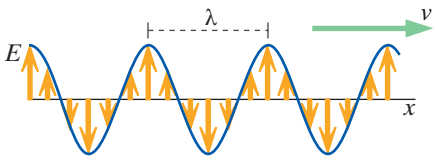


Figure S3.4 Space-varying electric field at one particular time.

Variation in Space

On the other hand, if you take a snapshot at one particular time, you see a sinusoidal or cosinusoidal pattern of electric field varying in space as shown in Figure S3.4.

This pattern in space can be expressed as $E\cos(2\pi x/\lambda)$, since the total phase of the cosine, $(2\pi x/\lambda)$, increases by 2π rad (360°) whenever x increases by an amount λ. Notice that if we replace x by t, and λ by T, this graph could just as easily represent the time variation of electric field at a particular fixed location in space.

Time and Space

A cosine wave that moves to the right can be described in terms of x and t as follows:

$$E\cos\left(2\pi\frac{t}{T}-2\pi\frac{x}{\lambda}\right)$$

This expression captures both the time and space aspects, because as time t increases, you need a larger x (motion to the right) to get the same total phase $(2\pi t/T-2\pi x/\lambda)$ corresponding to a particular crest of the wave.

> **QUESTION** In what direction is a wave traveling that is represented by the expression $E\cos(2\pi t/T+2\pi x/\lambda)$? (Note the + sign.)

In this case a larger t requires a smaller x in order to get the same total phase, so this expression represents a wave moving to the left.

Phase Shift

Depending on how we define $t=0$ or $x=0$, a wave may not be described exactly in the form $E\cos(2\pi t/T-2\pi x/\lambda)$ because it may not start at its maximum value. We can add a constant phase "shift" ϕ and write $E\cos(2\pi t/T-2\pi x/\lambda+\phi)$ to account for this. The phase shift ϕ can have any value between 0 and 2π (360°). When the phase shift between two waves is 0, we say the waves are *in phase*. When the phase shift is not zero, we describe the waves as *out of phase*.

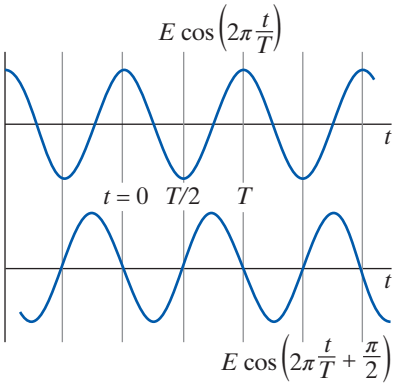


Figure S3.5 A phase shift affects the time at which the maximum occurs. The bottom curve differs from the top curve by a phase shift of $\pi/2$.

> **Checkpoint 1** Consider two examples of cosine waves in time, at the location $x = 0$ (Figure S3.5). Study the two curves, and make sure you understand the relationship between them.
>
> Label an x axis in units of half-periods ($t = T/2, 2T/2, 3T/2$, etc.). Then sketch a graph of $E\cos(2\pi t/T+\pi)$ to make sure you understand what a phase shift means. Your curve can be considered as either a phase shift by π radians or the negative of the original curve. Both points of view are useful.

Angular Frequency

Most of the time in our study of interference we only need the time variation of a wave at a particular location, in which case a cosine wave can be represented by any of these equivalent forms:

$E\cos(2\pi t/T+\phi)$	basic form
$E\cos(2\pi ft+\phi)$	since $f=1/T$
$E\cos(\omega t+\phi)$	where $\omega=2\pi f=2\pi/T$

The angular frequency ω (familiar from our study of orbits and vibrations in classical mechanics) is measured in radians per second (2π rad/T s). Because of its compactness, we will usually write $E\cos(\omega t+\phi)$.

Intensity

The frequency of electromagnetic waves in the visible range is so high that we don't perceive individual maxima and minima. Rather, we perceive a kind of time average. This time average is called the "intensity," and it is what we can most easily perceive and measure for many kinds of waves.

In Chapter 23 we saw that the energy flow carried by electromagnetic radiation is proportional to the square of the electric field. At a particular instant, the electric field is $E\cos(\omega t)$ and the instantaneous energy flow is proportional to $[E\cos(\omega t)]^2$. Therefore the time-averaged energy flow is proportional to the square of the amplitude (E^2).

We call the average energy flow the intensity I, measured in watts per square meter (W/m^2; see Figures S3.6 and S3.7). Although the electric field at a particular location varies sinusoidally, and the field is zero twice during every cycle, in many cases of interest the oscillation is so fast that we just perceive an average energy flow. For example, one cycle of visible light takes only about 1×10^{-15} s and your eye cannot detect the sinusoidal variation of such high-frequency oscillations.

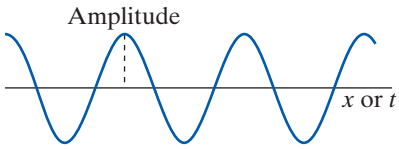


Figure S3.6 The amplitude is the maximum absolute value of the electric field in an electromagnetic wave.

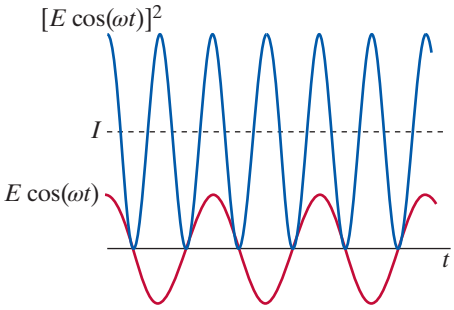


Figure S3.7 The time-averaged energy flow is called the intensity I.

INTENSITY AND AMPLITUDE

Intensity I (W/m^2) is proportional to (amplitude)^2.
$I \propto E^2$ for electromagnetic radiation.

Intensity is proportional to the square of the amplitude not only for sinusoidal electromagnetic radiation, but also for other kinds of sinusoidal waves, such as water waves and sound waves. In the case of visible light, the intensity is of course also called the "brightness." For sound waves, intensity corresponds to loudness. It is the intensity I, the time-averaged energy flow, that we typically measure.

Checkpoint 2 If the peak electric field in a sinusoidal electromagnetic wave is tripled, by what factor does the intensity change?

Interference of Two Electromagnetic Waves

According to the superposition principle, the net electric field at a location in space is just the vector sum of the electric fields contributed by *all* sources of electric field. This principle is still valid even if the electric fields change with time, so we can apply the superposition principle to electromagnetic radiation made by several different sources.

In Chapter 23 we saw that a radio antenna driven by sinusoidally alternating voltage (AC) produces propagating electric and magnetic fields whose magnitudes also vary sinusoidally. We begin our study of multiple sources of radiation by considering the simple case of just two such sources emitting sinusoidal radiation. Many situations involve very large numbers of sources, such as when light hits a mirror, but a careful analysis of the case of just two sources brings out the main aspects of the phenomena quite clearly.

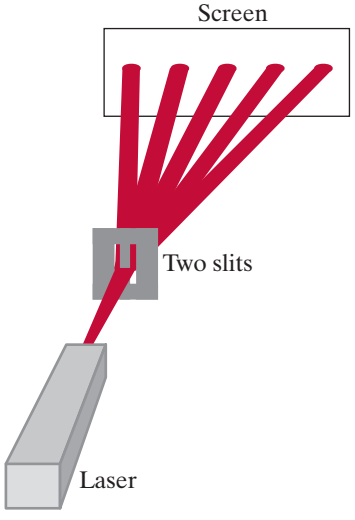


Figure S3.8 Laser light illuminates two narrow slits and makes an interference pattern on a screen.

Interference Pattern

When a laser illuminates two narrow vertical slits in a metal sheet, a pattern of multiple spots of light can be observed on a distant screen (Figure S3.8). Furthermore, the brightness of the spots decreases as the distance from the center of the pattern increases (this change in brightness is not shown in Figure S3.8).

QUESTION Does a particle model of light predict this pattern?

No. A simple particle model of light predicts that we will see only two bright spots, one in front of each slit. However, a wave model of electromagnetic radiation allows us to explain both the number of bright spots and their relative brightness. Furthermore, it allows us to make quantitative predictions of the locations of the bright and dark spots on the screen, simply by applying the superposition principle.

Coherent Monochromatic Radiation

Lasers emit light with rather special properties. The electromagnetic radiation in a laser beam is *coherent* (the phase of the oscillations is the same) and *monochromatic* (all radiation emitted has the same wavelength). Because a laser produces coherent, monochromatic light, it is a good source to use for experiments with light. In contrast, light from an ordinary incandescent light bulb is a mixture of radiation of many wavelengths, and is not coherent (electric fields not oscillating in phase).

We can think of the two slits as two separate sources of electromagnetic radiation, producing two electromagnetic waves that are coherent and monochromatic. A derivation at the end of this supplement shows that for the region between the slits and the screen we can treat the two-slit interference phenomenon as though the two slits were in-phase sources of light.

Constructive Interference

Consider two parallel radio antennas placed near each other, both driven by the same transmitter and both emitting electric-field cosine waves of the form $E\cos(\omega t)$, with the same amplitude E and the same angular frequency ω (where $\omega = 2\pi f = 2\pi/T$). For clarity, the accompanying magnetic field is not shown in Figure S3.9, nor do we show the $1/r$ fall-off of the radiative electric field. Since the emitted electromagnetic waves have only a single frequency, they are monochromatic. Because the phase of the waves emitted by the two sources is the same, they are also coherent.

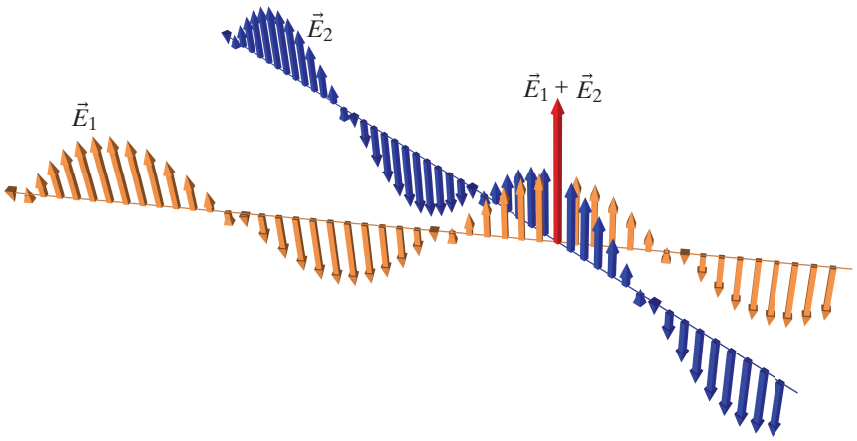


Figure S3.9 The electric field radiated by two radio antennas. At the particular location shown, the two electric fields are in phase and the net field is twice the field of one antenna.

At any location in space the net electric field is of course the superposition (vector sum) of the two radiative electric fields generated by the two antennas. At the location in Figure S3.9 where we show two waves crossing, the two electric fields add up in phase with each other:

$$E\cos(\omega t) + E\cos(\omega t) = 2E\cos(\omega t)$$

The maximum electric field (or amplitude) at this particular location is twice as large as the electric field due to a single antenna. We say that at this location the two waves interfere constructively.

Redistribution of Energy Flow

We have just seen that at a location where two cosine waves add up in phase, we find an amplitude that is twice as big as the amplitude due to a single antenna. (The electric field at this location still oscillates sinusoidally, with the same frequency as that of the two waves.)

> **QUESTION** If the amplitude is twice as big, how much bigger is the intensity (the time-averaged energy flow)?

The intensity at that location (or the brightness in the case of visible light) is four times as large as it would be from one antenna.

Evidently this provides a way to make more intense radiation at that location, but something is peculiar. The two antennas radiate twice as much energy as one antenna, so how can the intensity at this location be four times as large? Evidently there must be some other place where the intensity is smaller than usual.

Destructive Interference

Consider a different location, where the two cosine waves are out of phase by 180° (Figure S3.10). At this location, one of the waves had to travel $\lambda/2$ farther than the other one. Since they started in phase, there is now a phase difference of 180° (π rad).

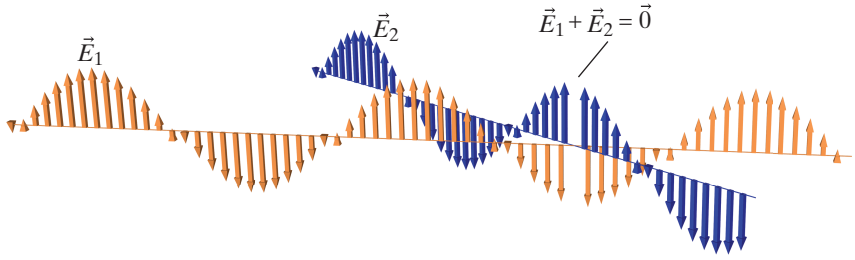


Figure S3.10 At a different location, the two electric fields are 180° out of phase and give a net field of zero at all times.

> **QUESTION** What is the amplitude (maximum electric field) at the location in Figure S3.10 where we show two rays crossing?

Evidently the electric field sums to zero at all times at this location, because the two waves are 180° out of phase with each other. We say that these two waves interfere destructively at this location, and the general phenomenon is called "interference." In physics the term interference refers in general to the superposition of waves from multiple sources. We say that at a location where the two waves add to each other, they interfere constructively, which is rather peculiar language but commonly used in technical discussions.

Two sinusoidal sources produce a complex pattern of radiation in space. The intensity of the radiation is larger at some locations than one might expect, and smaller or even zero at other locations. If you look at the intensity in every direction, it can be shown that the total energy radiated away is twice the energy radiated away by one antenna, as you would expect. It's just that the radiated energy is redistributed in a special way.

INTERFERENCE MAXIMA AND MINIMA

Two cosines can add up to give a *maximum* with twice the amplitude and four times the intensity. This occurs when the path difference is 0, λ, 2λ, 3λ, etc.

Two cosines can add to give a *minimum* with zero amplitude and zero intensity. This occurs when the path difference is $\lambda/2$, $3\lambda/2$, $5\lambda/2$, etc.

QUESTION Can you observe completely destructive interference (that is, zero intensity) at some location with two sources that have different angular frequencies ω_1 and ω_2? Why or why not?

There is no location at which the two sources cancel each other all the time, because even if the waves canceled each other at some moment, after one cycle of the first wave, the second wave wouldn't have returned to the same value, so they wouldn't cancel at this later time. Not only must the frequencies (or angular frequencies) of the sources be exactly the same, but the sources must also have a constant relative phase (must be coherent).

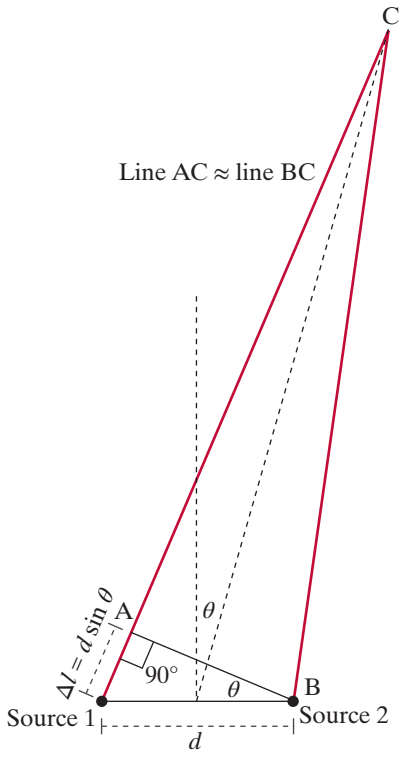


Figure S3.11 The path difference between two sources radiating in phase. Near the sources the paths are nearly parallel.

Predicting the Radiation Pattern for Two Sources

We should now be able to predict where two sources that are in phase with each other will give high-intensity maxima. All we have to do is figure out where the path differences are multiples of one wavelength. Equivalently, we can figure out where the phase difference is a multiple of 360° (2π rad). Similarly, we can find a location of zero intensity by figuring out where there is a path difference of half a wavelength or a phase difference of 180° (π rad).

Consider two sources that radiate in phase, and consider interference at a location C in a direction θ measured from the perpendicular bisector between the two sources (Figure S3.11). We need to be able to calculate the path difference Δl in terms of the angle θ, in order to determine whether there could be a maximum or a minimum in the intensity at location C.

Approximation for Distant Observation Locations

In Figure S3.11, if the observation location C is far away from the sources (as it usually is in most applications of the theory), we can make the approximation that in the right triangle ABC drawn on the diagram the base AC is almost the same length as the hypotenuse BC. This approximation is valid if the observation location C is far enough from the sources that $\angle$ACB is a very small angle. In this case, the path from source 1 to C is to an excellent approximation $\Delta l = d\sin\theta$ longer than the path from source 2 (path BC).

PATH DIFFERENCE

$$\Delta l = d\sin\theta$$

This is geometry, not physics. The physics lies in the fact that if the path difference Δl is an integer number of wavelengths, there will be a maximum at the angle θ determined by $\Delta l = d\sin\theta$. If Δl is $\lambda/2$, $3\lambda/2$, $5\lambda 2$, and so on, there will be a minimum (zero intensity) at the angle θ determined by $\Delta l = d\sin\theta$.

Note the importance of the spacing d between the sources. In particular, if $d < \lambda$ you can never get completely constructive interference except at $\theta = 0°$ or $\theta = 180°$, because $d\sin\theta$ is always less than λ. If $d < \lambda/2$ you can never get zero intensity (completely destructive interference).

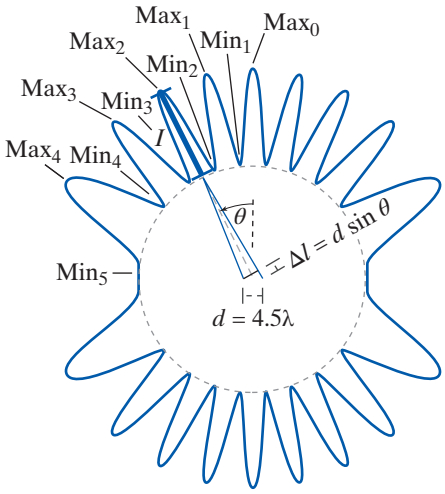


Figure S3.12 The intensity pattern for two sources. The bar marks the intensity I at the location of a maximum (Max_2).

Intensity vs. Angle

Figure S3.12 is a plot of the intensity pattern vs. angle for two sources that are driven in phase with each other and located a distance d apart, where in this particular case $d = 4.5\lambda$. In some directions (Max_0, Max_1, etc.), the difference in path length $\Delta l = d\sin\theta$ for the two cosine waves is zero or an integer number of wavelengths, the amplitudes add up in phase, and the intensity I is four times as large as the intensity of a single source. In some other directions (Min_1, Min_2, etc.), the difference in path length $\Delta l = d\sin\theta$ for the two cosine waves is half a wavelength (or $3\lambda/2$, $5\lambda/2$, etc.), the waves are 180° out of phase with each other, and the intensity I is zero.

> **QUESTION** Why is the intensity a maximum at $\theta = 0°$ (location Max_0) and $\theta = 180°$?

At 0° and 180° the path lengths are the same from both sources, so the waves are in phase, giving a maximum intensity.

> **QUESTION** In Figure S3.12, $d = 4.5\lambda$. Why is the intensity equal to zero at $\theta = 90°$ (location Min_5) and at $\theta = -90°$?

At 90° and −90° the difference in path lengths from the two sources is 4.5λ, so the two waves are 0.5λ out of phase, which is 180° difference in phase. This gives completely destructive interference.

Starting from $\theta = 0$, as θ increases from zero, the intensity falls rapidly to zero at location Min_1. The exact variation of the intensity with angle can be calculated by adding up two cosine waves with a varying phase difference, and this has been done in the computer program used to create the diagram. However, for many purposes it is sufficient just to figure out where the maxima and minima are. A rough qualitative sketch of the intensity pattern between a maximum and a minimum will be adequate for our purposes in this discussion.

Checkpoint 3 The two sources in Figure S3.12 are separated by a distance $d = 4.5\lambda$. **(a)** Calculate the angle θ where the first minimum is encountered (location Min_1 in Figure S3.12). (*Hint:* What must be the path difference $d\sin\theta$ in this direction?) **(b)** As θ increases beyond the location of the first minimum, the intensity rises rapidly to a second maximum, then falls to a minimum, rises again, and so on. Calculate the angle θ at location Max_4 in Figure S3.12. (*Hint:* What must $d\sin\theta$ be in this direction?)

Summary: Two-Source Interference

Here is a summary of the major points so far:

LOCATIONS OF MAXIMA AND MINIMA

Maximum intensity ($4I_0$) where $\Delta l = d\sin\theta$ is 0, λ, 2λ, 3λ, etc.
Minimum intensity (zero) where $\Delta l = d\sin\theta$ is $\lambda/2$, $3\lambda/2$, $5\lambda/2$, etc.
No zero-intensity minimum is possible if $d < \lambda/2$.

If the sources are coherent but *not* in phase, you need to take that phase difference into consideration in addition to the phase difference due to Δl.

Two-Slit Interference

Earlier we mentioned the phenomenon of two-slit interference, observable with a laser light source. We are now in a position to relate the slit spacing, angles at which we find maxima and minima, and the wavelength of the laser light.

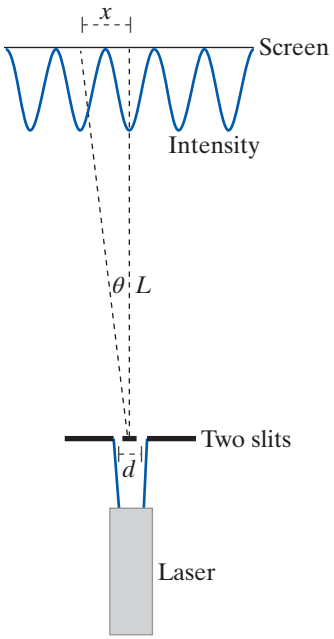


Figure S3.13 Top view of a laser, two slits, and an interference pattern on a screen.

In a top view (Figure S3.13), we show the laser, the two slits, and the intensity pattern on the screen, which can be measured quantitatively with a light meter. Suppose that the distance between the slits is $d = 0.5$ mm, the distance from the slits to the screen is $L = 2$ m, and the distance on the screen from the center maximum to the next maximum is measured to be $x = 2.4$ mm.

QUESTION Calculate the wavelength of this laser light.

At the second maximum, $d\sin\theta = \lambda$, so $\sin\theta = \lambda/d$. For small angles, $\sin\theta \approx \tan\theta \approx \theta$, so $\lambda/d \approx x/L$. Therefore we have

$$\lambda = \frac{xd}{L} = \frac{(2.4\times10^{-3}\text{ m})(0.5\times10^{-3}\text{ m})}{(2\text{ m})} = 6\times10^{-7}\text{ m} = 600\text{ nm}$$

This wavelength is in the red portion of the visible spectrum. It was experiments of a similar kind that initially established what we now know about the wavelengths of visible light.

Intensity Variation of Interference Maxima

We can predict the locations of interference maxima and minima. However, we have not yet found an explanation for the possible variation in brightness of the maxima. To explain why in some cases the central maximum may be much brighter than the others, we will need to consider diffraction from a single slit, which is discussed in Section S3.3.

Wavelength, Path Length, and Distance between Sources

The interference effects we have described show up in situations where the difference in path length of two rays is larger than but comparable to the wavelength of the electromagnetic radiation. In the case of radio waves we considered in the previous section, d, the spacing between the antennas, was several times the wavelength, resulting in path differences of up to 4λ. In the case of two-slit interference with a laser, the slit spacing (0.5 mm, or 5×10^{-4} m) was nearly 1000 times the wavelength of the light (about 500 nm, or 5×10^{-7} m). However, because the surface on which we saw the interference pattern was quite far from the slits, the difference in path length for the two waves was still only a few times the wavelength of the light.

Interferometry

A modern application of this interference phenomenon is to measure very small distances. A laser beam is split into two beams, and mirrors guide one of the beams to an object of interest, from which the beam is reflected to a screen where there is interference with the other laser beam. If the object moves even a fraction of one wavelength of visible light, there is an easily observable shift in the interference pattern on the screen. This use of interference phenomena to measure small changes in distance is called "interferometry."

S3.2 MULTISOURCE INTERFERENCE: DIFFRACTION

So far we've discussed interference effects produced by two coherent sources. "Diffraction" is the term used to describe interference effects involving a large number of sources; "interference" describes effects involving only a few sources. In this section we'll discuss several examples of diffraction.

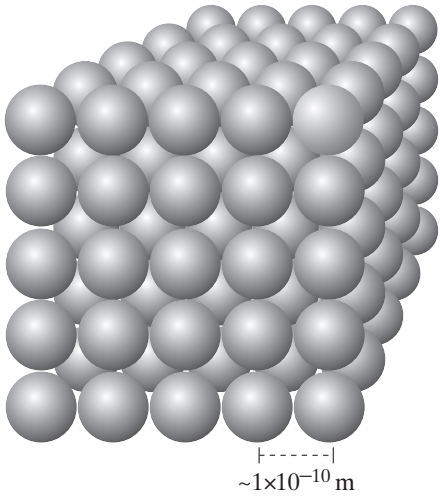


Figure S3.14 Part of a simple cubic crystal lattice. Other more complex geometric arrangements of atoms are also possible.

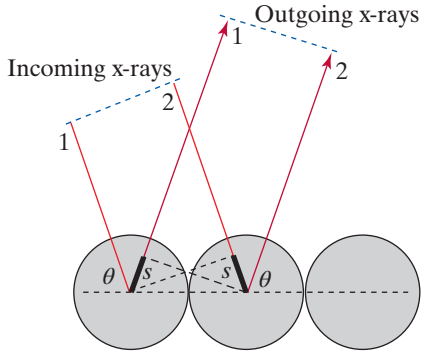


Figure S3.15 Re-radiation from the top layer of atoms in a crystalline solid. Note that θ is measured from the surface, not from the normal to the surface.

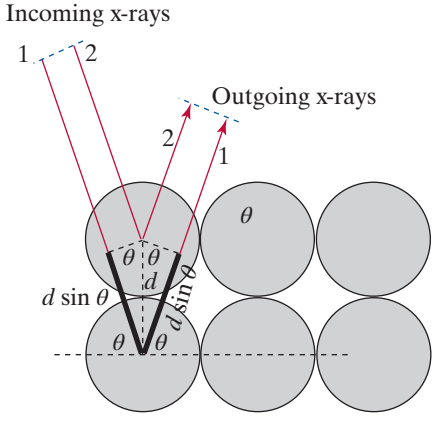


Figure S3.16 Re-radiation from the top two layers involves an additional path difference.

X-Ray Diffraction

Light in the x-ray region of the spectrum can interact with the electrons in an atom in a solid crystalline object. In order to probe the structure of a crystal, it is necessary to produce a monochromatic, coherent beam of x-ray radiation, and to find a way to measure precisely the wavelength of this radiation. X-ray diffraction, a technique that can be used to study the structure of crystals or of individual molecules in a crystal, can also be used to prepare a monochromatic x-ray beam and to measure precisely the wavelength of such a beam. Because of its importance, we will discuss x-ray diffraction in some detail.

In a crystal, atoms or molecules are arranged in a regular three-dimensional array, called a *lattice,* as shown in Figure S3.14. If electromagnetic radiation of a single wavelength hits a crystal, electrons in all the atoms are accelerated and re-radiate. We can expect to find high intensity re-radiation only in those directions where the path difference for adjacent atoms is λ, 2λ, 3λ, and so on. Dealing with a three-dimensional array of sources can be geometrically complex, but we will limit this discussion to some simple cases that illustrate the basic principle.

X-Ray Wavelength: $\lambda \approx d$

The spacing between atoms in a crystalline solid is of the order of magnitude of atomic sizes, or about 1×10^{-10} meter. We have seen that interference effects require that the wavelength λ must be comparable to the spacing d, or about 1×10^{-10} m. Such electromagnetic radiation is called x-rays. X-rays have very short wavelengths compared to visible light: violet light has a wavelength of 400 nm, or 4000×10^{-10} m, about 4000 times the wavelength of x-rays.

A common way to produce x-rays is to shoot a beam of very high-energy electrons into a block of metal, such as copper. These high-energy electrons sometimes knock an electron out of an inner electron shell of the copper atom. Very shortly thereafter an electron from an outer electron shell drops into the hole, and the associated large drop in energy of the atom is accompanied by the emission of light of high energy and short wavelength. This is the production mechanism used in medical x-ray equipment.

Conditions for Constructive Interference

To see the basic idea behind x-ray diffraction, we'll consider a beam of single-wavelength x-rays that hit a simple crystal, one in which the atoms are arranged on a three-dimensional rectangular grid. If the x-rays come in at an angle θ to the surface, the first layer of atoms will re-radiate with constructive interference at an angle θ that looks like simple reflection, because in that direction there is a zero path difference between adjacent sources (atoms in the first layer). Along path 1 in Figure S3.15, there is an extra length s in the outgoing ray, and along path 2, there is the same extra length s in the incoming ray. At other angles the path difference would not be zero, and unless the path difference is an integer number of wavelengths there would not be constructive interference. We will consider just the simple case where we observe a maximum at the simple reflection angle.

The re-radiation from other atomic layers, below the surface layer, need not be in phase with the top layer unless special conditions hold. In Figure S3.16, if the center-to-center distance between layers is d, there is a difference $2d\sin\theta$ in path lengths for electromagnetic radiation accelerating electrons in an atom in the top layer (which re-radiates), and for accelerating electrons in an atom in the next layer (which also re-radiates).

If the difference in path length, $2d\sin\theta$, is not equal to the wavelength λ (or some multiple of λ), the two topmost layers of atoms will not re-radiate in

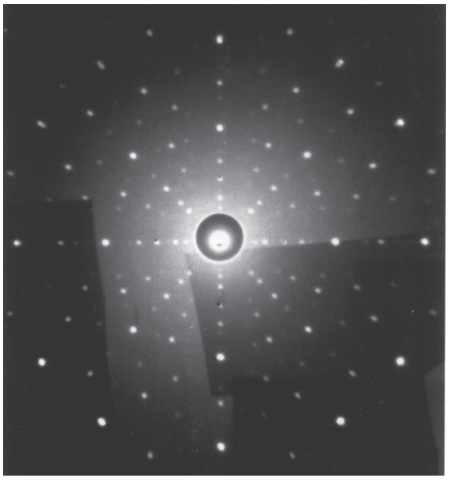


Figure S3.17 X-ray diffraction pattern showing interference maxima from a single crystal of tungsten. (Photo by Barry Luokkala, Department of Physics, Carnegie Mellon University)

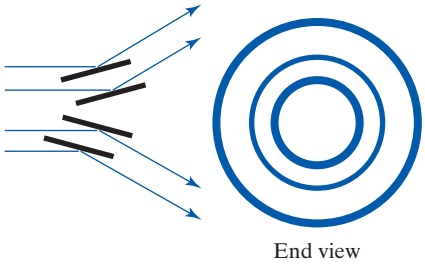


Figure S3.18 X-ray diffraction with a powder of crystals at random orientations.

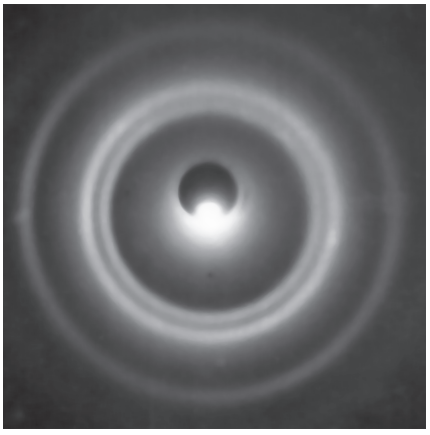


Figure S3.19 X-ray diffraction pattern from polycrystalline lithium fluoride. (Photo by Barry Luokkala, Department of Physics, Carnegie Mellon University)

phase at the angle θ. Also at some other angle θ, the atoms in the top layer won't radiate in phase with each other. Therefore if you orient the surface of the crystal at an angle θ to the x-ray beam, you may or may not get an intense beam of x-rays re-radiated (reflected) from the crystal at the reflection angle. You will get an intense "reflection" only if the following condition is true:

CONDITION FOR CONSTRUCTIVE X-RAY INTERFERENCE

$$2d\sin\theta = n\lambda, \quad \text{where } n \text{ is an integer}$$

Note that if the top two layers re-radiate in phase, the re-radiation from all the other layers will also be in phase.

This effect, called x-ray diffraction because it involves a huge number of sources, can be used to deduce the spacing d between layers in a crystal. Assume that you have a source of single-wavelength x-rays, whose wavelength λ is known. You direct a beam of these x-rays at an angle θ onto the crystal, and you look for an intense beam at the reflection angle. If there is no intense beam there, you change the angle a bit and look again. (Typically one turns the crystal rather than the large x-ray beam apparatus.) If you find an intense reflection at a particular angle θ, you know that the spacing between layers of atoms is $d = n\lambda/(2\sin\theta)$. You may need to do some additional checking to determine the value of n, which could be 1, 2, 3, and so on. Figure S3.17 shows a pattern of constructive interference from a crystal of tungsten.

The x-rays produced by some kinds of x-ray sources have a continuous spectrum of wavelengths. How can one make a beam of single-wavelength x-rays from such sources?

> **QUESTION** Suppose that you have a crystal with known spacing d between atomic layers. Explain briefly how to use this crystal with a multi-wavelength beam of x-rays to produce a single-wavelength beam, which can be used in experiments on crystals whose structure is not known.

For the given atomic-layer spacing d and desired wavelength λ, we can get a large intensity for just that wavelength at the angle θ that satisfies $2d\sin\theta = \lambda$. If we direct the x-ray beam at the crystal at this angle, we'll get a strong reflection for the desired wavelength, whereas other wavelengths won't give a strong beam in that direction.

X-ray diffraction is an extremely powerful tool that has been used to determine the structure of a huge variety of crystals. It can also be used with polycrystalline materials, materials in the form of a powder of tiny crystals oriented in random directions. In such a powder containing a very large number of tiny crystals, there will be some crystals that happen to be oriented at angles to the x-ray beam that satisfy the condition $2d\sin\theta = n\lambda$, in which case there is intense re-radiation in a ring surrounding the incoming x-ray beam, at the appropriate angle (Figure S3.18). X-ray film exposed to this re-radiation shows rings whose diameters tell a great deal about the crystal structure, even if one doesn't have a single large crystal available. A diffraction pattern from powdered lithium fluoride is shown in Figure S3.19.

A layer of atoms re-radiating in phase need not be parallel to the surface of a crystal. A layer can consist of any geometrical plane drawn through the crystal in such a way that many atoms lie in that plane (so as to give a large re-radiation intensity). In addition, if the crystal is made up of more than one kind of atom (as in NaCl crystals, for example), accurate measurements of intensity can provide information about the arrangement of the different atoms, because different atoms re-radiate different amounts of energy.

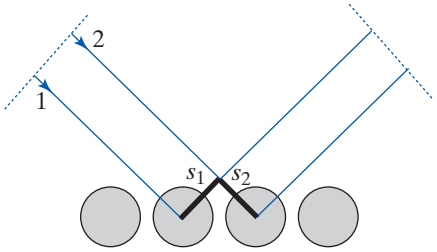


Figure S3.20 Only when incident angle = reflection angle is $s_1 = s_2$, so $\Delta l = 0$.

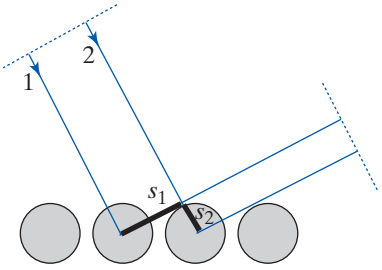


Figure S3.21 At any other angle, $s_1 \neq s_2$, so Δl is nonzero.

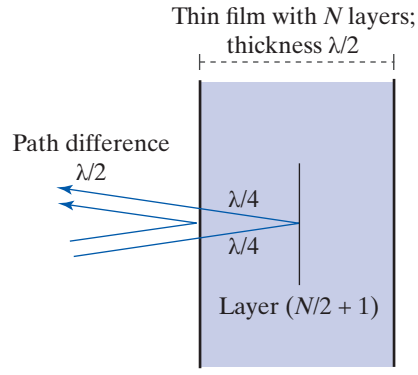


Figure S3.22 Visible light encounters a very thin film of matter. Consider pairs of layers such as layer 1 and layer $(N/2+1)$, etc.

Checkpoint 4 It is known that the spacing between neighboring planes of atoms in a particular crystal is 2×10^{-10} m. A monochromatic x-ray beam of wavelength 0.96×10^{-10} m is directed at the crystal. At what angle might one expect to find an interference maximum?

Reflection of Visible Light: $\lambda \gg D$

Why do we see a reflection of visible light from a smooth surface at only one angle? When visible light instead of x-rays hits a crystal, it accelerates electrons in the material, and there is indeed re-radiation in all directions. However, the x-ray diffraction condition $2d\sin\theta = n\lambda$ for adjacent layers is not relevant for visible light, because the distance between atomic layers is extremely small compared to the wavelength of visible light ($d \approx 1 \times 10^{-10}$ m, $\lambda \approx 5000 \times 10^{-10}$ m). The difference in path length for re-radiation from adjacent layers is negligible, and adjacent layers are nearly in phase.

Because the interatomic spacing d is much smaller than λ for visible light, there will be constructive interference between two waves only if the difference in path length is zero, so we will see only one high-intensity maximum. As can be seen from Figure S3.20, for interactions with atoms in the surface layer $\Delta l = 0$ only if the reflection angle equals the incident angle. At any other angle, the difference in path length is nonzero (Figure S3.21), and there is destructive interference. The interaction of the incident radiation with deeper layers of atoms is essentially the same, since the layers are very close together compared to the wavelength of the radiation.

Metal Surfaces

Reflection of visible light off a metal surface is quite different from the reflection off transparent insulating materials such as water or glass. The mobile electrons in a metal are very easy to accelerate. The mobile electrons near the surface gain a lot of energy, so we conclude that little energy is available to transfer to electrons far from the surface. It must be the case that the intensity of the radiation (and therefore the intensity of the electric field) inside the metal is small. How can this happen? Presumably there is so much re-radiation by accelerated electrons near the surface of the metal that there is very little net field in the interior of the metal. There are no significant interference effects between layers at different depths from the surface. Of course there are the usual interference effects involving electrons in different locations on the surface, so that there is strong intensity only at the "reflection" angle, and metals make good mirrors.

Thin-Film Interference

Despite what we said about the condition $2d\sin\theta = n\lambda$ for adjacent layers in connection with reflection of visible light, there can be interesting interference effects when visible light hits a thin film of transparent material only a few wavelengths thick, such as a soap bubble, or a thin oil slick floating on a puddle of water. These thin-film interference effects depend critically on the exact thickness of the thin film.

Assume for simplicity that single-frequency incoming light hits the surface of a thin film nearly head-on (rays nearly perpendicular to the surface, as in Figure S3.22). Suppose that the thin film has a thickness of only half a wavelength $(\lambda/2)$. Though this is a very thin film, it nevertheless contains a large number N of atomic layers: N is about $(250 \times 10^{-9}\ \text{m}/1 \times 10^{-10}\ \text{m})$ =2500 atomic layers.

In Figure S3.22, consider pairing the first layer of atoms with atomic layer number $(N/2+1 = 1251)$, pairing the second atomic layer with atomic layer

number ($N/2 + 2 = 1252$), and so on. Atomic layer number ($N/2 = 1250$) is paired with atomic layer number $N = 2500$ (the last layer). The path difference for each of these pairs is $\lambda/2$, so each pair produces zero intensity, and the total re-radiated intensity is zero for a thin film whose thickness is $\lambda/2$.

By similar arguments you can show that there will be zero re-radiation for film thicknesses of $\lambda/2$, $2\lambda/2$, $3\lambda/2$, $4\lambda/2$, and so on—any integer multiple of $\lambda/2$. For film thicknesses halfway between these thicknesses ($\lambda/4$, $3\lambda/4$, $5\lambda/4$, etc.), it can be shown that there is maximum re-radiation, although a formal proof is beyond the scope of this discussion.

QUESTION A soap film or oil slick when illuminated by ordinary white light shows a rainbow of colors. Briefly explain why this is.

For a given film thickness, some wavelength might be the right length to give fully constructive interference and be bright, but other wavelengths wouldn't be the right length and would be dim. If the film thickness isn't uniform or you look at a different angle, the wavelength that gives fully constructive interference will be different.

In a soap bubble that lasts a long time, more and more of the liquid flows to the bottom of the bubble. The top part gets very thin, thinner than a quarter-wavelength of light even for violet light, which has the shortest visible wavelength. This very thin film seems to re-radiate almost no light at all (it appears black), because there are only a small number of atomic layers compared to the several thousand that contribute when the thickness is a quarter-wavelength.

A closely related thin-film interference phenomenon is responsible for the brilliant iridescent colors seen in some butterfly wings and bird feathers, which contain structures with spacing comparable to a wavelength of light. There is constructive interference for some colors and destructive interference for others.

Index of Refraction

The relevant wavelength in thin-film interference is the wavelength inside the material. As we saw in Chapter 23, this turns out to be shorter than the wavelength in air. Inside a dense transparent material such as glass or water, individual atoms are significantly affected not only by the incoming electromagnetic radiation but also by the electric fields re-radiated by the other atoms. When you add up all the electric fields, you find that the pattern of the net field has a crest-to-crest distance that is shortened. The factor by which the wavelength is decreased is called the index of refraction and is around 1.5 for many kinds of glass; it is 1.33 for water.

In a material with index of refraction n, the wavelength $\lambda' = \lambda/n$, where λ is the wavelength of the radiation in vacuum. The frequency of the wave is not affected, and since $v = f\lambda$ the *apparent* speed of light inside a solid is noticeably slower than 3×10^8 m/s.

This complication does not affect our analysis of x-ray diffraction because the re-radiation by an atom exposed to x-rays is very small compared to the re-radiation by an atom exposed to visible light. In the case of x-rays, we can consider the electric field inside the material to be due almost entirely to the incoming radiation, with a negligible contribution from the tiny re-radiation from the other atoms.

Checkpoint 5 A film of oil 200 nm thick floats on water. The index of refraction is 1.6. For what wavelength of light will there be complete destructive interference?

Coherence Length

There is an effect that limits interference effects in films. Most sources of visible light produce sinusoidal waves with a fairly short total length L, which corresponds to the short length of time $\Delta t = L/c$ during which an atom was emitting the light (c is the speed of light). Even if the total length L corresponds to a large number of wavelengths, it might be less than a millimeter. If the total length L is shorter than twice the thickness of the film, by the time the light re-radiated by the last atomic layer emerges from the film, the first atomic layer has stopped emitting, and there is nothing to interfere with. The total length L of the sinusoidal emission is called the "coherence length." One of the important properties of lasers is that they emit light with a very long coherence length. This is particularly important in making holograms, which are based on interference effects.

Diffraction Gratings

A compact disc (CD) looks pretty much like a mirror, and you can see a lamp reflected in the shiny surface. However, if you hold the disc in a fixed position and move your eye away from the direct reflection you see a rainbow of colors. This is an interference effect. The compact disc has concentric rings of tiny pits etched into the plastic, and these rings of pits separate undisturbed rings of plastic from each other. You observe interference between reflections from adjacent undisturbed sections of the plastic. The distance d between adjacent rings is comparable to the wavelength of visible light (otherwise there would be no effect). There are a very large number of rings on a compact disc, and the effect is called "diffraction," the term applied to interference phenomena when there are a large number of sources whose waves interfere with each other.

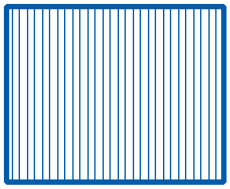


Figure S3.23 A diffraction grating (not to scale) has many fine parallel scratches on its surface.

A similar device is the diffraction grating, a piece of glass or plastic that has been accurately scratched with a large number of closely spaced straight parallel lines a very small distance d apart (Figure S3.23). As with a compact disc, you can see an ordinary reflection in a diffraction grating, but off at an angle you see a rainbow of colors. Transmission diffraction gratings are transparent between the scratches and act like a large number of parallel slits through which light passes. As with two-slit interference, each slit can be treated as though it were a source of light (see the derivation at the end of this supplement).

The diffraction grating is an important scientific tool because it can be used to measure very accurately the wavelengths (and therefore the frequencies) of light emitted by atoms. Different atoms emit light of different frequencies, and diffraction gratings have not only played an important role in the study of atomic structure but have also provided a convenient way to identify the atoms present in a sample of material.

A diffraction grating may have as many as 1000 parallel scratches per millimeter! In this case $d = 1 \times 10^{-3}$ mm $= 1 \times 10^{-6}$ m; compare with a typical wavelength of light of about 500 nm $= 0.5 \times 10^{-6}$ m.

> **QUESTION** Before we treat this phenomenon quantitatively, see whether you can explain very briefly and qualitatively why there are bands of colors off at an angle to a compact disc (or reflective diffraction grating), and why the ordinary direct reflection does not show a rainbow.

The angle θ at which you get another maximum depends on the wavelength (because the phase difference between adjacent strips depends on the wavelength). Therefore, different wavelengths in the white light show up at different angles. The ordinary reflection is normal, because in this case the path lengths are all equal, so there is no dependence on wavelength.

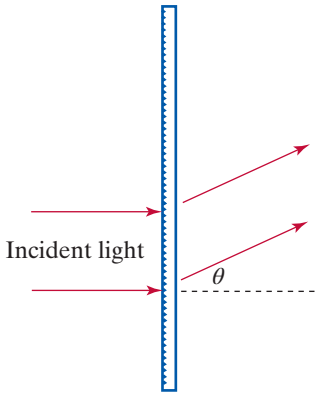


Figure S3.24 A transmission diffraction grating.

To analyze a diffraction grating quantitatively, we'll consider a particularly simple geometry but one that illustrates the most important aspects of the device. Suppose that a beam of red light strikes a transmission diffraction grating nearly perpendicular to the grating (Figure S3.24). There will be a central peak at $\theta = 0$ since all the slits are in phase with each other, but there can be one or more additional peaks in other directions, where the path difference between adjacent slits is one wavelength.

> **QUESTION** If the spacing between adjacent slits is d and the wavelength of the red light is λ, what is the first nonzero angle θ at which you see high intensity? (*Hint:* Think through what θ is required to get a path difference of one wavelength for adjacent slits of the grating; draw a diagram for adjacent slits.)

This maximum occurs at an angle where $d \sin\theta = \lambda$, so the angle at which you will see high intensity for red light ($\lambda \approx$ about 700 nm) is calculable from $\sin\theta = \lambda/d$.

> **QUESTION** If you replace the source of red light with a source of violet light (wavelength of about 400 nm), will the high intensity be seen at a larger or smaller angle θ? Why?

Since $\sin\theta = \lambda/d$, a shorter wavelength λ will give a maximum at a smaller angle θ. Since you see these various colors coming at different angles from the diffraction grating (or compact disc), this implies that white light from the Sun or from an incandescent lamp must be a mixture of many colors, from red to yellow to green to violet.

Diffraction gratings are used heavily in science and technology to determine what wavelengths are present in various kinds of light, which is an indicator of what atoms are producing that light. If you look through a diffraction grating at a slit illuminated by a neon lamp, you can see individual lines of color that are characteristic of the neon atoms inside the lamp that are producing the light. This line spectrum is different for different atoms, and quite different from the continuous spectrum produced by hot glowing objects such as the Sun or an incandescent lamp.

What's Important about Having Lots of Slits?

In a later section on angular resolution we will show formally that the beams of light at the maxima of a diffraction grating are extremely narrow compared to those made by just two slits. It is easy to see from an energy argument that this must be true. With just two slits, the maximum intensity is $2^2 = 4$ times the intensity due to one slit. However, in a typical diffraction grating 2 cm wide, with spacing between slits $d = 1 \times 10^{-3}$ mm, there are 2×10^4 slits, and the maximum intensity is $(2 \times 10^4)^2 = 4 \times 10^8$ times the intensity of a single slit! For there to be such a large intensity in a few special directions, there must be wide regions where the intensity is extremely small. The beams of light from a diffraction grating are extremely narrow and sharply defined, which is why diffraction gratings can be used to make very precise measurements of wavelengths.

Checkpoint 6 A certain diffraction grating is known to have a line spacing of $d = 1 \times 10^{-3}$ mm. A source of single-wavelength light is observed to produce a high-intensity beam at an angle $\theta = 36^\circ$ to the normal. If this is the first maximum at a nonzero angle, what is the wavelength of the light?

This wavelength is emitted by excited sodium, as in sodium vapor lamps. Detecting this wavelength in some light indicates the presence of excited sodium.

EXAMPLE **Experimenting with a CD**

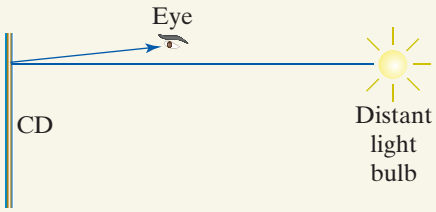


Figure S3.25 Looking straight at a CD, you see an ordinary reflection.

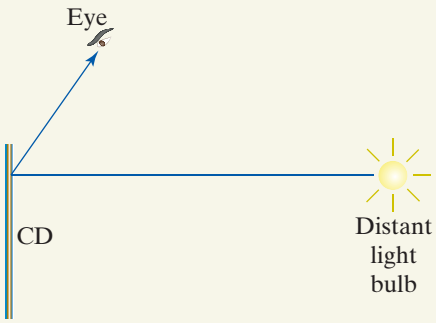


Figure S3.26 Looking at an angle to a CD, you see colors.

Hold a compact disc (CD) nearly perpendicular to a line from the CD to a distant light bulb. If you look nearly perpendicular into a region of the CD near the edge, you see an ordinary reflection of the light bulb (Figure S3.25). While holding the CD stationary in this position, move your head slowly away from the perpendicular (or equivalently, tip the CD) while continuing to look at this region of the CD, until you see the first color of a rainbow (Figure S3.26).

(a) Is the first color you see red or violet? (Try it!) Why? (Violet ≈ 400 nm; red ≈ 700 nm.)

(b) If you keep moving your head further away from the perpendicular (or tipping the CD) you see a complete spectrum. If you continue to even larger angles you may see additional spectra. How many complete spectra do you see?

(c) Use your observations in parts (a) and (b) to estimate the distance between circular tracks on the CD. Report the data on which you base your estimate.

(d) The CD is read with a relatively inexpensive infrared laser (wavelength longer than red). Why can't the tracks be placed closer together, to get more information on the CD? (Newer devices such as DVD players use shorter-wavelength lasers.)

Solution First, try the experiment for yourself!

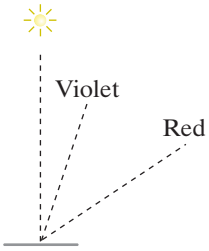


Figure S3.27 Violet and blue appear at the smallest angle. Yellow and red appear at larger angles.

(a) Moving your head away from the perpendicular, the first color you see is violet, which has the shortest wavelength of the visible spectrum (Figure S3.27). This is what you would expect, because for the light re-radiated by all the grooves to be in phase, $\sin\theta = \lambda/d$, so the smallest wavelength will have a maximum at the smallest angle (d is the spacing between grooves on the CD).

(b) Most people see three spectra, but this is a bit tricky, since you have to hold your head just right.

(c) We found the first maximum for red light to be at an angle θ_1 of about 30°, and we can obtain a value for d from this:

$$d\sin\theta = \lambda \quad \text{(first maximum)}$$

$$d = \frac{\lambda}{\sin\theta_1} \approx \frac{(700\times10^{-9}\text{ m})}{\sin 30^\circ} = 1400\times10^{-9}\text{ m} = 1400\text{ nm}$$

Alternatively, if we see three complete spectra, the red of the third spectra is an at angle $\theta_2 \approx 90^\circ$, so we can carry out an alternative, approximate calculation of the spacing d.

$$d\sin\theta = 3\lambda \quad \text{(third maximum)}$$

$$d = \frac{3\lambda}{\sin\theta_1} \approx \frac{3(700\times10^{-9}\text{ m})}{\sin 90^\circ} = 2100\times10^{-9}\text{ m} = 2100\text{ nm}$$

(d) If the spacing d between the tracks is shorter than the wavelength of the incident light, it is not possible to resolve the images of the two tracks (they overlap). If the tracks were closer together, the wavelength of light used to read the information would need to be shorter.

Scattering of Light

In Chapter 23 we discussed why the sky is blue, which involves frequency-dependent re-radiation by air molecules of light from the Sun,

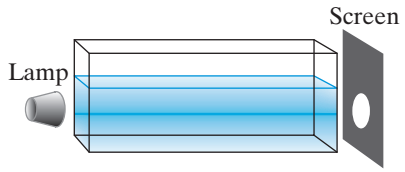


Figure S3.28 Light passes through a tank of water and appears on a screen.

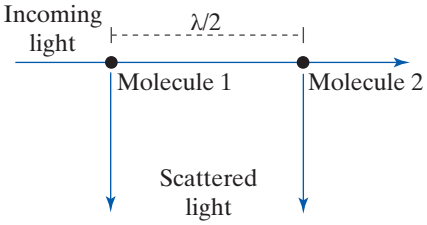


Figure S3.29 Viewed from above: re-radiation (scattering) from water molecules a distance $\lambda/2$ apart.

with blue light re-radiated more strongly than red light. This process, in which light passing through a nearly transparent medium is partially re-radiated to the side, is called "scattering" (the light is scattered into directions different from the original direction).

We will describe an experiment that brings out some interference aspects that affect scattering phenomena. Perhaps your instructor will demonstrate this for you. A bright white light is directed through a long aquarium tank full of water and onto a screen, where it shows up as a white spot (Figure S3.28). If the water is quite clean, very little light is scattered to the side.

A puzzle is why there is almost no light scattered off to the side. After all, there are huge numbers of molecules in the water, and light striking those molecules accelerates electrons, leading to re-radiation. We'll consider path differences for different molecules and see where we should expect low-intensity and high-intensity light.

For any molecule in the water, find a second molecule a short distance $\lambda/2$ downstream (Figure S3.29). Off to the side, the path difference for the re-radiation from these two molecules is half a wavelength, so the intensity is zero. If we can pair up all the molecules in this way, we can expect to see very little intensity off to the side.

Can we always find a suitable partner for every molecule? The answer for water (or glass) is yes, because in a liquid or solid the molecules are right next to each other, filling the entire space. Go downstream a distance $\lambda/2$, a distance of about 250 nm = 2500 atomic diameters, and you can be sure that you will find a molecule at that location (unless you are within $\lambda/2$ of the end of the water tank, but that short region contains very little water). It is true that molecules in a liquid have some limited freedom to move about, unlike molecules in a solid, but they do fill the space at all times. Note that in the downstream direction, straight through the water toward the screen, the re-radiation from these two molecules is in phase, so we can expect to see light coming out the end of the water tank, and we do.

(It is reasonable to ask what would happen if we paired up molecules that were a distance λ apart. Wouldn't that give constructive interference? Yes, but we would also have to calculate how that re-radiated light would interfere with the light from all the other pairs that were one wavelength apart. This calculation is beyond the scope of this discussion, but the result is that we get zero intensity when we add up all the contributions. The simplicity of the argument in terms of pairs of molecules that are $\lambda/2$ apart comes from the fact that it is easy to add up all the zero intensities from all such pairs and get zero.)

Add Some Soap to the Water

Next we drop some material such as soap into the water to provide a low concentration of specks of material. Now you see some light scattered off to the side, and it is distinctly bluish. At the same time, the spot on the screen becomes somewhat reddish, because the blue portion of the white light has been scattered more than the red portion, leaving light that has been depleted of the blue component. As explained in Chapter 23, the scattered bluish light is strongly polarized, as you can see by turning a piece of polarizing material in front of your eye as you look toward the tank from the side. (You also get a big effect by turning a piece of polarizing plastic between the light source and the water tank.)

The visible scattering is due to there being a relatively small number of specks of material in the water that re-radiate a different amount than the water molecules they displace. Moreover, the specks are far away from each other in random locations and do not destructively interfere with each other. That is, if you go a distance $\lambda/2$ downstream from a soap particle, you won't necessarily find another, similar soap particle at that location.

Scattering from a Gas

There is a subtle point concerning why the sky is blue. Why is there significant scattering off to the side of clear air, when we see little scattering from water? A gas is different from a liquid or solid in that the molecules are not in contact with each other but are roaming around quite freely, with lots of space between molecules. Since the air molecules are not in relatively fixed positions, the argument for destructive interference that we used for water is not valid.

Consider a cube of air whose edge is only 0.1λ long, at standard temperature and pressure. Such a cube contains a surprisingly large number of molecules, though many fewer than in a comparable volume of a liquid or solid:

$$N = (0.1 \times 500 \times 10^{-9}\ \text{m})\frac{6 \times 10^{23}\text{molecules}}{22.4\ \text{L}}\frac{1\ \text{L}}{1 \times 10^{3}\ \text{cm}^3}\frac{1 \times 10^{6}\ \text{cm}^3}{\text{m}^3}$$
$$\approx 3000\ \text{molecules}$$

If there were exactly 3000 molecules in each tiny cube of air, there would be little scattering, due to destructive interference by radiation from pairs of cubes located a distance of $\lambda/2$ apart. However, the number of molecules in each cube fluctuates randomly as air molecules roam in and out of a cube. The scattering from the sky that we see and enjoy is largely due to the fluctuations in the number of molecules per unit volume. This effect is much less pronounced for liquids. Although the molecules do slide past each other in a liquid, the number of molecules per unit volume hardly changes, since the molecules remain in contact with each other. Of course there is no fluctuation at all in a solid, and if there were no impurities or spatial imperfections in a block of glass it would not scatter light passing through it.

The typical statistical fluctuation of N molecules in a given volume of air can be shown to be equal to the square root of N. The electric fields re-radiated by these $\sqrt{N}$ molecules are nearly in phase with each other (because they are in a volume that is small compared to λ). Thus, the intensity is proportional to $(\sqrt{N})^2 = N$. The scattering is what one would naively expect—N times the scattering from one molecule!

S3.3 ANGULAR RESOLUTION

In this section we will examine a set of phenomena that at first may seem unrelated but can be explained by the same basic aspects of the wave model of light. We are interested in the following phenomena:

- The diameter of a telescope mirror (or lens) determines whether you see two distant stars as two separate bright spots or one large bright spot.
- In two-slit interference, not all the maxima have the same brightness.
- Light passing through one single slit shows maxima and minima on a screen placed in front of the slit.
- The width of a diffraction grating determines the angular separation of the maxima observed through the grating.

All of these phenomena involve what is called "angular resolution," and can be explained in terms of the equation for angular spread

$$\Delta\theta = \frac{\lambda}{W}$$

where W is the width of the device.

Angular Width of a Maximum

An important property of a diffraction grating is that in the direction of a maximum, the re-radiated beam is extremely narrow. Just how wide (in angle)

is one of the maxima observed when monochromatic light passes through a particular diffraction grating? The narrower this maximum, the easier it is to distinguish light of two slightly different wavelengths.

Width of a Maximum

How do we define the "width" of a maximum? We could say that the width of the maximum is the angular distance from the brightest area of the maximum to the darkest area next to it—the adjacent minimum. Our task is to calculate the location of the adjacent minimum.

Suppose that θ is the angle of the first maximum, with the path difference for adjacent slits being one wavelength, so that

$$d\sin\theta = \lambda$$

Since the maximum is at angle θ, we'll call the location of the adjacent minimum $(\theta + \Delta\theta)$. Our task is to find the value of $\Delta\theta$.

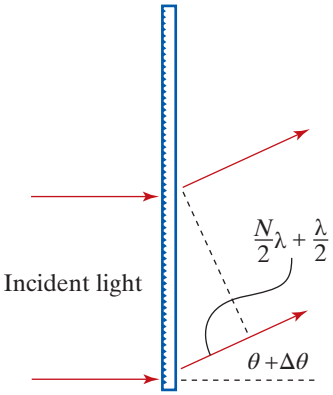


Figure S3.30 The first slit and the middle slit of the diffraction grating interfere destructively if the path difference for adjacent slits is $\lambda + \lambda/N$.

Conditions for a Minimum

We note that there will be a minimum if light passing through the first slit and the $(N/2)$th slit (the middle slit) interferes destructively (Figure S3.30). This is reminiscent of the argument we made when considering sideways scattering of light from a liquid. Assume that there are an even number N of slits (or if N is odd, neglect the small amount of light from the last slit).

For the first nontrivial maximum (not straight ahead), the path difference between adjacent slits is one wavelength, so the path difference between the first slit and the $(N/2)$th slit is $(N/2)\lambda$. For a minimum, this path difference has an additional $\lambda/2$:

$$\left(\frac{N}{2}\right)\lambda + \frac{1}{2}\lambda$$

Now consider the next pair of slits, slit number 2 and slit number $(N/2+1)$. They also have a path difference of $\lambda/2$, so they too interfere destructively. We have succeeded in pairing up all the slits in such a way that they each contribute zero intensity, so the total intensity in the direction $\theta + \Delta\theta$ is zero.

Write the path difference between adjacent slits like this:

$$d\sin(\theta + \Delta\theta) = \lambda + \varepsilon \quad \text{where } \varepsilon \text{ is a small fraction of } \lambda$$

The path difference between the first slit and the $(N/2)$th slit is this:

$$\frac{N}{2}(\lambda + \varepsilon) = \left(\frac{N}{2}\right)\lambda + \frac{1}{2}\lambda$$

Solving for ε we find

$$\varepsilon = \frac{\lambda}{N}$$

Finding $\Delta\theta$

Now, to find $\Delta\theta$ we use trig identities to expand this relation for two adjacent slits:

$$d\sin(\theta + \Delta\theta) = \lambda + \varepsilon = \lambda + \frac{\lambda}{N}$$

$$d\sin\theta\cos\Delta\theta + d\cos\theta\sin\Delta\theta = \lambda + \frac{\lambda}{N}$$

Since $\Delta\theta$ is a small angle, we can make the approximations that

$$\cos\Delta\theta \approx 1 \quad \text{and} \quad \sin\Delta\theta \approx \Delta\theta$$

so

$$d\sin\theta + d\cos\theta\Delta\theta \approx \lambda + \frac{\lambda}{N}$$

Because $d\sin\theta = \lambda$

$$d\cos\theta\Delta\theta \approx \frac{\lambda}{N} \quad \text{and} \quad \Delta\theta \approx \frac{\lambda}{Nd\cos\theta}$$

The total width of the diffraction grating W is N times the width of one slit:

$$W = Nd$$

Unless θ is close to 90°, $\cos\theta$ is on the order of 1, so we find that $\Delta\theta$, the "half width" of the maximum, is approximately given by this:

ANGULAR HALF WIDTH OF A MAXIMUM

$$\Delta\theta \approx \frac{\lambda}{W}$$

where W is the total width of the (illuminated portion) of the device.

Although we've proved this only for a diffraction grating, this result is quite general. Whenever a device contains a very large number of sources that interfere, the angular width of a maximum-intensity beam is approximately equal to the wavelength divided by the total width of the device. For visible light hitting a diffraction grating that is 2 cm wide, the angular spreading may be very small: $\Delta\theta \approx \lambda/W \approx (600\times10^{-9}\text{ m})/(2\times10^{-2}\text{ m}) = 3\times10^{-5}$ rad, which is less than 2 millidegrees.

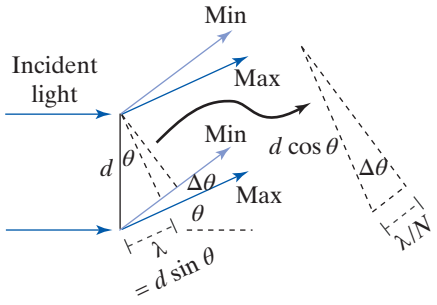


Figure S3.31 Two adjacent slits in a transmission diffraction grating, with a maximum at θ and a minimum at $\theta + \Delta\theta$.

An Alternative Geometric Argument

Instead of using trig identities as we did above, we can find $\Delta\theta$ by a geometric argument involving two adjacent slits (Figure S3.31). In the dashed triangle, the side opposite the very small angle $\Delta\theta$ is $\varepsilon = \lambda/N$, because this is the additional path length that yields a minimum. The hypotenuse is approximately $d\cos\theta$. The angle $\Delta\theta$ in radians is approximately equal to the ratio of the opposite side to the hypotenuse:

$$\Delta\theta \approx \frac{\lambda/N}{d\cos\theta} = \frac{\lambda}{Nd\cos\theta} = \frac{\lambda}{W\cos\theta}$$

where W is the total width of the grating ($W = Nd$). This is an extremely important result: The wider the grating, the narrower the beam. Unless θ is close to 90°, $\cos\theta$ is of the order of 1, so we get, as before:

$$\Delta\theta \approx \frac{\lambda}{W}$$

What about Other Pairings of Slits?

You might wonder whether we would not have gotten complete cancellation if we had chosen other ways of pairing up the slits, rather than pairing the first slit with the middle slit, and so on. For example, the first and last slits have a path difference of $N\lambda + \lambda$, so they interfere constructively. However, this is the only pair that does this, and if we were to add up the effects of all the other slits we would find that their combined effects would cancel the re-radiation of the outer slits. Once we have shown that there is some pairing that gives a net intensity of zero, we don't actually have to consider other pairings.

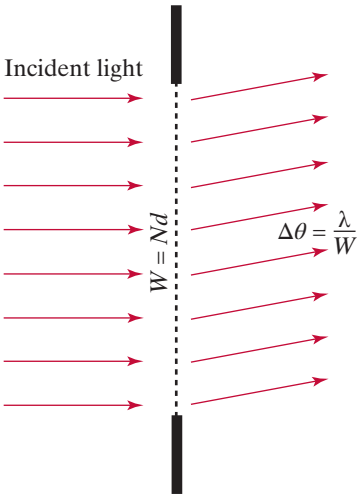


Figure S3.32 The first minimum for a single slit is at $\Delta\theta \approx \lambda/W$.

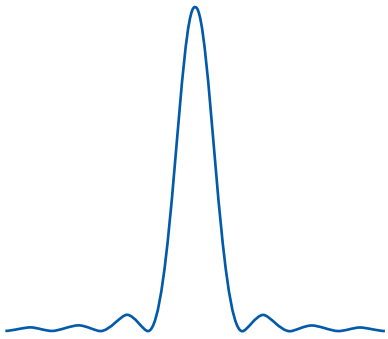


Figure S3.33 The pattern of intensity vs. angle for diffraction by a single slit.

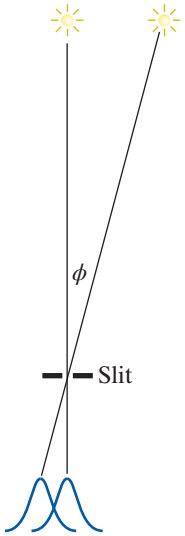


Figure S3.34 Light from two distant stars passes through the aperture of a telescope. The half width of the diffraction-broadened image for each star determines whether or not it will be possible to distinguish the two images.

Single-Slit Diffraction

What does a diffraction grating have to do with maxima and minima created when light passes through a single slit? Just as two narrow slits through which light passes are equivalent to two narrow sources of light, so we can model a single wide slit, of width W, to be equivalent to a very large number N of very narrow sources located a distance $d = W/N$ apart. (Actually, the sum over N finite sources turns into an integral over an infinite number of infinitesimal slits.) There is of course a maximum straight ahead, but there is a minimum at a small angle $\Delta\theta$ given approximately by λ/W (Figure S3.32). The overall pattern of intensity looks like Figure S3.33.

This also explains the variation in intensity of the maxima in two-slit interference. Each slit makes a pattern like Figure S3.33, and these two patterns interfere. The intensity pattern is the combined result of both single-slit diffraction and two-slit interference.

Checkpoint 7 To get a feel for the size of these "diffraction" effects, consider a single slit $W = 0.1$ mm wide, illuminated by violet light. The wavelength λ of violet light is about 400 nm (400×10^{-9} m). Light from the slit hits a screen placed a distance $L = 1$ m $= 1000$ mm away from the slit, and the image of the slit is much larger than 0.1 mm due to significant amounts of light heading at angles as large as plus or minus λ/W (for a total angular spread of $2\lambda/W$). Calculate this total width in millimeters.

Telescope Resolution Limited by Diffraction

When a telescope makes an image of a distant star on a photographic plate or array of electronic detectors, the image is broadened by diffraction, which spreads the light rays out into a cone with an approximate half-angle $\Delta\theta \approx \lambda/W$. The images of two stars that are close together in the sky may be broadened so much that they overlap each other and cannot be "resolved" into two separate images (Figure S3.34). The "resolving power" or "resolution" of a telescope or camera is limited by diffraction. The wider the lens, the sharper the image (bigger W, smaller $\Delta\theta$). Alternatively, the shorter the wavelength, the sharper the image (smaller λ, smaller $\Delta\theta$). One rule of thumb for resolution states that if ϕ, the angle between the stars, is greater than $2\Delta\theta$, the images will be resolvable.

Radio telescopes used in making astronomical observations are sometimes linked electronically to other radio telescopes located thousands of miles away, in order to make a radio-frequency "lens" that is so large that the diffraction broadening is extremely small. This has made it possible to obtain exquisitely detailed pictures of extremely distant structures.

It is much more difficult to do this with ordinary optical telescopes that capture visible light, because you have to bring the light from two different telescopes to the same place, where there can be interference. Nevertheless, there are now linked pairs of large optical telescopes with mirrors to bring light from the two telescopes to an observation location between the telescopes. This makes it possible to obtain pictures with much higher resolution than can be obtained with older optical telescopes.

Only One Large Diffraction Peak

With a few sources we often find several large completely constructive peaks, corresponding to the path difference between neighboring sources being equal to $0, \lambda, 2\lambda$, and so on. Even a diffraction grating with many sources may produce several maximum-intensity beams. The situation is different with single-slit diffraction.

QUESTION Explain briefly why a single slit (or lens) produces just *one* large peak (there are smaller peaks, but only one large peak).

With a single slit, the path difference d between adjacent sources is infinitesimal in size, so the path difference is an infinitesimal quantity and can never get as big as λ. There is no direction (except for $\theta = 0$) where all the amplitudes are in phase.

Checkpoint 8 Satellites used to map the Earth photographically are typically about 100 mi (160 km) above the surface of the Earth, where there is almost no atmosphere to affect the orbit. If the lens in a satellite camera has a diameter of 15 cm, roughly how far apart do two objects on the Earth's surface have to be in order that they can be resolved into two objects, if the resolution is limited mainly by diffraction? The key point is that the angle subtended by the two objects should be comparable to or bigger than the spread in angle $2\Delta\theta$ introduced by diffraction of light going through the lens. Visible light has wavelengths from 400 to 700 nm; consider a wavelength somewhere in the middle of the spectrum. You should find that the numbers on a car license plate would be too blurred to read.

S3.4 MECHANICAL WAVES

Mechanical waves have many features in common with waves of light. An example of a mechanical wave is the propagation of sound in a solid, which was discussed in Chapter 4. Other examples of mechanical waves include waves running along a taut string, water waves, and sound propagating through air.

In the following pages we'll derive the wave equation, a general equation that describes all kinds of waves:

$$\frac{\partial^2 u(x,t)}{\partial t^2} = v^2 \frac{\partial^2 u(x,t)}{\partial x^2}$$

In this partial differential equation u represents some quantity that is "waving" in time or space or both, with the wave moving along the x axis. For example, the quantity u can represent the deflection of an atom in a solid through which sound is propagating away from the atom's equilibrium position, or the sideways deflection of a piece of string along which a pulse is moving, and v is a constant equal to the speed of propagation. (The quantity u could also represent the electric field in electromagnetic radiation at some location in space, in which case $v = c$.)

The meaning of $\partial u(x,t)/\partial t$ is "the derivative of u with respect to t, holding x constant." This partial derivative represents the rate of change with time of the deflection of an atom (or of the electric field) at some particular location x. Similarly, the meaning of $\partial u(x,t)/\partial x$ is "the derivative of u with respect to x, holding t constant." That is, take a snapshot of the wave at some particular time t and see how u changes from one location x to a neighboring location.

The Wave Equation: Longitudinal Waves

In Chapter 4 we discussed the speed of sound in a solid and presented a dimensional analysis that suggests the speed ought to be approximately $v \approx \omega d = \sqrt{k_{s,i}/m_a}\, d$, where $k_{s,i}$ is the effective stiffness of the interatomic bond modeled as a spring, and m_a is the mass of one atom. We will derive the wave equation for this situation, starting from a microscopic view, and use it to show that our expression for v is indeed correct as long as the wavelength is much longer than the interatomic distance d, which is almost always the case.

Figure S3.35 shows three of the atoms in a long row of atoms that are connected by springs representing the interatomic bonds. We choose our coordinate system so that the first atom at the far left is located at $x = 0$ (not

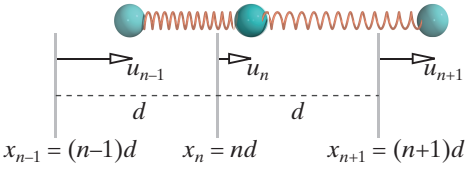


Figure S3.35 Three of the atoms in a long row of atoms connected by springs representing the interatomic bonds. Each atom has been displaced from its equilibrium position by an amount u.

shown), so that the equilibrium position of the nth atom is nd if the equilibrium distance between atoms is d. If a pulse travels along this row of atoms, the atoms will be temporarily displaced from their normal, equilibrium positions. This is called a longitudinal wave in which the displacements of the atoms are in the same direction as the wave motion.

We'll call the instantaneous displacement of the nth atom from its equilibrium position u_n, where the instantaneous position of the nth atom is $nd + u_n$ as shown in the snapshot in Figure S3.35. We'll calculate the net force acting on the nth atom in terms of the stiffness of the springs, and then use the Momentum Principle to determine how the momentum of the nth atom changes with time. We'll show that the Momentum Principle takes the form of the wave equation, from which we'll be able to determine the speed of propagation of a pulse along the row of atoms (the speed of sound).

For generality, the row of atoms may be under tension (or compression), so that the distance d between atoms may be larger (or smaller) than d_0, the relaxed length of a spring.

The Force Acting on the nth Atom

We'll calculate the x component of the net force on the nth atom, which for convenience we'll write briefly as F_n. We will calculate how much the springs to the right and left of the nth atom are stretched (or compressed), which we can calculate in terms of the instantaneous positions of the atoms and the unstretched length of the spring, which for a solid neither in tension or compression is just the equilibrium interatomic distance d (see Figure S3.35):

$$\begin{aligned}
\text{stretch to the right} &= L - L_0 = [((n+1)d + u_{n+1}) - (nd + u_n)] - d \\
&= u_{n+1} - u_n \\
\text{stretch to the left} &= L - L_0 = [(nd + u_n) - ((n-1)d + u_{n-1})] - d \\
&= u_n - u_{n-1} \\
F_n &= k_{s,i}[\text{stretch to the right} - \text{stretch to the left}] \\
&= k_{s,i}[(u_{n+1} - u_n) - (u_n - u_{n-1})]
\end{aligned}$$

Notice that in the expression for F_n the interatomic spacing d cancels out. However, we'll see that d does come into the speed of sound we'll calculate.

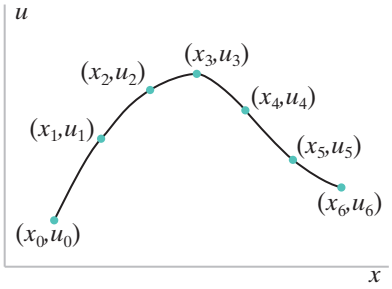


Figure S3.36 We imagine a continuous function $u(x)$ that passes through the actual atomic positions. This is a snapshot of the wave at a particular time.

It is useful to think of u as a smooth (continuous) function of x. This is illustrated in Figure S3.36, a snapshot of the wave at a particular time, where we imagine a continuous function $u(x)$ that passes through the actual positions of the atoms. In our expression for the force on the nth atom the term $(u_{n+1} - u_n)$ shows up in the definition of the derivative $\partial u/\partial x$ (holding t constant) if we approximate a small Δx by the equilibrium distance between atoms, d:

$$\begin{aligned}
\frac{\partial u}{\partial x} &= \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} \\
&\approx \frac{u_{n+1} - u_n}{d}
\end{aligned}$$

We can usefully take this idea one step further by considering the second derivative of u:

$$\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) &= \lim_{\Delta x \to 0} \frac{\Delta(\partial u/\partial x)}{\Delta x} \\
&\approx \frac{(u_{n+1} - u_n)/d - (u_n - u_{n-1})/d}{d} \\
\frac{\partial^2 u}{\partial x^2} &\approx \frac{(u_{n+1} - u_n) - (u_n - u_{n-1})}{d^2}
\end{aligned}$$

The numerator in this expression can be found in the expression for the force on an atom, so we can write the following, where we remind ourselves that u is actually a function of both x and t:

$$F \approx k_{s,i} d^2 \frac{\partial^2 u(x,t)}{\partial x^2}$$

QUESTION What approximations did we make in deriving this expression for the force on an atom?

Our result depends on the validity of treating u as a continuous function of x that can be differentiated, whereas in reality u has values only at locations $x = 0$, d, $2d$, etc. We can expect our result to be quite accurate as long as the shape of the traveling pulse changes relatively slowly from one atom to the next, which is typically the case for sound waves traveling through a solid. For example, middle C on a piano has a frequency of 256 Hz (256 cycles per second). In aluminum the speed of sound is $v = 4800$ m/s, so the wavelength of a 256 Hz sound wave in aluminum is (4800 m/s)/(256 cycles/s) = 18.75 m. This wavelength is huge compared to the distance between atoms, so u in this case is indeed a function that changes slowly from one atom to the next.

A reflection: Why is the force proportional to a second derivative? We found that the stretch to the right of the nth atom is $u_{n+1} - u_n + d$ and the stretch to the left is $u_n - u_{n-1} + d$. As we have seen, $u_{n+1} - u_n$ and $u_n - u_{n-1}$ are related to the first derivative of u with respect to x. If both springs are stretched by the same amount, the net force on the nth atom is zero. This corresponds to a row of atoms (or a macroscopic wire made up of these atoms) which is in tension due to forces applied at the ends, but there is no wave motion, only equilibrium. For there to be wave motion, somewhere along the row of atoms there has to be a difference between the stretch to the right and the stretch to the left, which means a difference in $\partial u/\partial x$, which means a nonzero second derivative $\partial^2 u/\partial x^2$. An atom may be in motion where the net force on the atom is zero, though its velocity would not be changing at that instant.

Applying the Momentum Principle

The rate of change of the momentum of the nth atom is of course equal to the net force acting on this atom, and from applying the Momentum Principle to this atom we will be able to determine the speed of a pulse running along the row of atoms. For speeds small compared to the speed of light

$$\frac{dp_x}{dt} \approx m_a \frac{dv_x}{dt} = m_a \frac{d^2 x}{dt^2} = F_{\text{net}_x}$$

where m_a is the mass of one atom. The position of the nth atom is $x_n = nd + u_n$. The rate of change of the momentum of the nth atom in the chain of atoms is

$$m_a \frac{\partial^2 (nd + u_n)}{\partial t^2} = m_a \frac{\partial^2 u_n}{\partial t^2}$$

since n and d do not change with time. Note that $d\vec{p}/dt$ involves time, and $\vec{F}$ involves positions.

Dropping the subscript n, the Momentum Principle for any atom is the following, where again m_a is the mass of each atom and $k_{s,i}$ is the effective stiffness of the interatomic spring:

$$m_a \frac{\partial^2 u(x,t)}{\partial t^2} = F = k_{s,i} d^2 \frac{\partial^2 u(x,t)}{\partial x^2}$$

$$\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{k_{s,i} d^2}{m_a} \frac{\partial^2 u(x,t)}{\partial x^2}$$

This has the form of the wave equation introduced earlier if we make this replacement:

$$v^2 = \frac{k_{s,i}d^2}{m_a}$$

$$v = \pm\sqrt{\frac{k_{s,i}}{m_a}}\,d$$

We will show that $|v|$ is the speed of propagation of the wave, and the sign determines the direction of propagation. This is the same result we obtained from dimensional analysis in Chapter 4 for the speed of sound in a solid, but just from dimensional analysis alone we wouldn't know whether the speed might be twice this value, or one-third this value, etc. The plus-minus sign for v, resulting from taking the square root, means that either a wave moving to the right or a wave moving to the left is a solution of the wave equation.

> **QUESTION** Suppose you stretch a row of N atoms, so that the interatomic distance increases to $d+s$. How does this affect the speed of sound v? (*Hint:* Examine how the net force depends on s.) How does this affect the time for a wave to propagate from one end of the row of atoms to the other end?

If you examine our calculation of the net force on an atom, you'll find that s in the stretch to the right and s in the stretch to the left cancel each other, so the speed of sound v doesn't change. The time it takes for a wave to move the length of a row of N atoms, a total length which is now $N(d+s)$, increases because the wave has (slightly) farther to go. The change in the length of a metal bar when you stretch it is of course typically quite small, so this is a small effect.

A Solution of the Wave Equation

Despite its apparent complexity, the wave equation has a rather simple solution. We can show that the function

$$u(x,t) = f(x-vt)$$

satisfies the wave equation and represents a pulse traveling to the right (the $+x$ direction) along the chain of atoms, where v is the speed of the wave, no matter what the exact form of the function f is. For example, $u(x,t) = 5\cos(10(x-3t))$ is a possible solution of the wave equation. A component of the electric field in a propagating electromagnetic wave could be expressed as $E(x,t) = E_0 \sin k(x-ct)$.

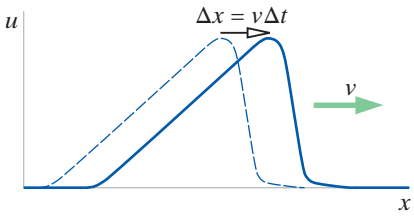


Figure S3.37 A pulse travels to the right along a row of atoms a distance $\Delta x = v\Delta t$ at speed v for a time Δt. The quantity u represents the deflection of an atom away from its equilibrium position.

In Figure S3.37, at time t the peak of a pulse is at location x and at a slightly later time $t+\Delta t$ the peak of the pulse has moved to the nearby location $x+\Delta x$, where $\Delta x = v\Delta t$. For the peak of the wave to have moved as shown in Figure S3.37 we must have

$$f(x+\Delta x - v(t+\Delta t)) = f(x-vt)$$

This will be true if $\Delta x - v\Delta t = 0$, which means that $v = \Delta x/\Delta t$, which is of course just the definition of the speed. Evidently a pulse described by $u(x,t) = f(x-vt)$, where v is a positive number, represents a pulse moving to the right with speed v. Another way to see that $u = f(w)$ must represent a wave traveling to the right, where $w = (x-vt)$, is that for the function argument $w = (x-vt)$ to stay the same as the time t increases, the location x must increase.

> **QUESTION** What kind of a wave is represented by the function $u(x,t) = f(x+vt)$, with v a positive number?

As time t advances (increases), x must decrease for the function argument $(x+vt)$ to remain the same. Therefore $u(x,t)=f(x+vt)$ must represent a wave traveling to the left.

Is such a traveling pulse consistent with the wave equation? We need to take partial derivatives of the function $u(w)$, where $w=(x-vt)$. Note that the derivative of u with respect to w is an ordinary, not a partial derivative, because u is a function of the single variable $w=(x-vt)$.

$$\frac{\partial u}{\partial t}=\frac{du}{dw}\frac{\partial w}{\partial t}=\frac{du}{dw}(-v) \quad \text{using the differentiation chain rule}$$
$$\frac{\partial^2 u}{\partial t^2}=\frac{\partial}{\partial t}\left(\frac{\partial u}{\partial t}\right)=\frac{\partial}{\partial t}\left(\frac{du}{dw}(-v)\right)=\frac{d}{dw}\left(\frac{du}{dw}(-v)\right)\frac{\partial w}{\partial t}=v^2\frac{d^2u}{dw^2}$$
$$\frac{\partial u}{\partial x}=\frac{du}{dw}\frac{\partial w}{\partial x}=\frac{du}{dw}(1) \quad \text{using the differentiation chain rule}$$
$$\frac{\partial^2 u}{\partial x^2}=\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)=\frac{\partial}{\partial x}\left(\frac{du}{dw}(1)\right)=\frac{d}{dw}\left(\frac{du}{dw}\right)\frac{\partial w}{\partial x}=\frac{d^2u}{dw^2}$$

These results show that we can write the following, which shows that $u=f(x-vt)$ is in fact consistent with the wave equation:

$$\frac{\partial^2 u(x,t)}{\partial t^2}=v^2\frac{\partial^2 u(x,t)}{\partial x^2}$$

By showing that the v in the wave equation represents the speed of propagation of a wave, we have proved that the speed of sound in a solid is indeed $v=\sqrt{k_{s,i}/m_a}\,d$.

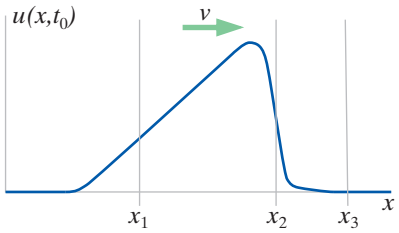


Figure S3.38 A snapshot at time $t=t_0$ of a pulse traveling to the right at speed v. What is the position as a function of time of the sections of rope at locations x_1, x_2, and x_3?

A Traveling Pulse

Figure S3.38 is a snapshot of a pulse propagating to the right along a taut rope. Being a snapshot, this shows $u(x,t_0)$ at all locations x, at one particular instant $t=t_0$. The meaning of $u(x,t)=f(x-vt)$ can be made clear by considering the future motions of the rope at the locations marked x_1, x_2, and x_3. Here are graphs of u vs. t at these three locations, starting at time $t=t_0$:

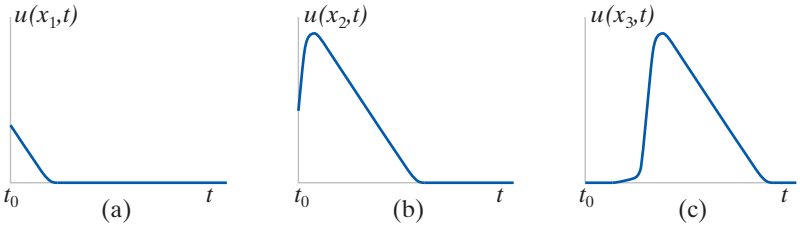


Figure S3.39 u as a function of t at the three different locations shown in Figure S3.38: x_1, x_2, and x_3.

Location x_1: As the pulse moves to the right, u decreases (see Figure S3.38 and Figure S3.39a); the velocity of this section of rope is downward. Eventually the trailing edge of the pulse passes location x_1, and this section of rope no longer moves.

Location x_2: As the pulse moves to the right, u at first increases (velocity upward), reaches a maximum as the peak of the pulse passes (velocity momentarily zero), then decreases (velocity downward; see Figure S3.38 and Figure S3.39b). Again, when the trailing edge of the pulse passes location x_2, this section of rope no longer moves.

Location x_3: When the leading edge of the pulse reaches location x_3, u begins to move upward. It reaches a maximum as the peak of the pulse passes location x_3, then decreases until the trailing edge of the pulse passes location x_3,

after which this section of the rope no longer moves. The graph of u vs. t (Figure S3.39c) looks like a mirror image of the pulse shape, u vs. x (Figure S3.38).

Note that the graphs of u vs. t in Figure S3.39 have a different horizontal scale than the graph of u vs. x in Figure S3.38.

A Macroscopic View of the Speed of Sound

In Chapter 4 we saw that Young's modulus Y, a macroscopic ratio of stress (force per unit area) to strain (fractional stretch) can be expressed in terms of microscopic quantities: $Y = k_{s,i}/d$. Also, the density ρ of the material can be written as $\rho = m_a/d^3$, because each atom occupies a tiny cube d on a side (in the simple case of a cubic lattice). The speed of sound in a solid can be expressed in terms of macroscopic quantities:

$$v = \sqrt{\frac{k_{s,i}}{m_a}}\, d = \sqrt{\frac{Yd}{m_a}}\, d = \sqrt{\frac{Yd^3}{m_a}} = \sqrt{\frac{Y}{\rho}}$$

It can be shown that for sound waves in air, which are longitudinal waves in which the air pressure changes in connection with changes in the density, $v = \sqrt{B/\rho}$, where the "bulk modulus" B is defined as the change in pressure P divided by the associated fractional change in density ρ:

$$B = \frac{\Delta P}{\Delta\rho/\rho} = \frac{1}{\rho}\frac{dP}{d\rho}$$

B has the same units as Young's modulus Y: force per unit area, N/m^2. Sound waves in air propagate at a speed of about 340 m/s at room temperature.

Superposition

There is a superposition property for solutions of the wave equation. If both functions $f(x,t)$ and $g(x,t)$ are solutions of the wave equation, then the sum of these two solutions, $h(x,t) = f(x,t) + g(x,t)$, is also a solution of the wave equation:

$$\frac{\partial^2(f(x,t)+g(x,t))}{\partial t^2} = v^2\frac{\partial^2(f(x,t)+g(x,t))}{\partial x^2}$$

$$\frac{\partial^2 f(x,t)}{\partial t^2} + \frac{\partial^2 g(x,t)}{\partial t^2} = v^2\frac{\partial^2 f(x,t)}{\partial x^2} + v^2\frac{\partial^2 g(x,t)}{\partial x^2}$$

The first terms on the left and right are equal, and the second terms on the left and right are equal, so the wave equation is satisfied. A simple corollary of this superposition property is that multiples of a solution such as $6.3\,f(x,t)$ are solutions if $f(x,t)$ is a solution.

This important property of superposition is true because the wave equation is "linear" in the solution function, which appears only to the first power. If the function appeared in the wave equation squared, say, then we wouldn't have been able to show that the sum of two solutions was also a solution.

This superposition property will be important in a later section on standing waves.

The Wave Equation: Transverse Waves

If the atoms move sideways, perpendicular to the direction of the wave motion, that is called a transverse wave. An example is a wave moving along a rope, or water waves. We'll analyze a transverse wave traveling along a row of atoms connected by springs representing the electric interatomic force. The analysis

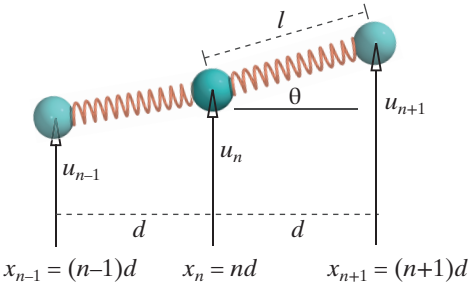


Figure S3.40 A snapshot of three neighboring atoms in a row of atoms along which a transverse wave is passing.

is very similar to the analysis of a longitudinal wave, but the details of the geometry lead to a different prediction for the wave speed.

In Figure S3.40 we show a snapshot of the positions of atoms as a transverse wave passes by. Similar to our analysis of longitudinal waves, we'll call the transverse position of the nth atom u_n.

Consider the spring to the right of the nth atom in Figure S3.40. We are going to make the approximation that the angle θ is small enough that the length l of the spring is nearly the same as d, the equilibrium distance between atoms. This small-angle approximation can be quite accurate. It corresponds to saying $d/l = \cos\theta \approx 1$. You can test this on your calculator: What is $\cos(5°)$? (0.996.) What is $\cos(10°)$? (0.985.) As long as we can assume that the angle of a spring to the horizontal is small, the length l of the spring is always approximately equal to d, the equilibrium distance between atoms.

Since the force that a spring exerts on an atom is proportional to the stretch, and all the springs have nearly the same length d, the magnitude of the force exerted by each spring is nearly the same, $k_{s,i}(d - d_0)$, where d_0 is the relaxed length of a spring. This is called the tension $F_T = k_{s,i}(d - d_0)$, which is nearly the same all along the chain of atoms.

Although the magnitude of each spring force is (nearly) the same, the directions are not the same, as you can see in Figure S3.40. For small angles $\sin\theta \approx \tan\theta \approx u_n/d$. You can test this on your calculator: What is $\sin(5°)$? (0.0872.) What is $\tan(5°)$? (0.0875.) Therefore the x and u components of the spring forces applied to the left and right of the nth atom are these:

$$
\begin{aligned}
F_{n,x} &= F_T\cos\theta_n - F_T\cos\theta_{n-1} \\
F_{n,x} &\approx F_T - F_T = 0 \\
F_{n,u} &= F_T\sin\theta_n - F_T\sin\theta_{n-1} \\
F_{n,u} &\approx F_T\left[\frac{u_{n+1}-u_n}{d} - \frac{u_n - u_{n-1}}{d}\right]
\end{aligned}
$$

As in the analysis of longitudinal waves, we imagine a continuous function $u(x,t)$ that passes through the points (x_n, u_n) and assume that the change in u from one atom to the next is small (long wavelength compared to the interatomic distance d). We take a snapshot (freeze the motion) and write partial derivatives with respect to x:

$$
\begin{aligned}
\left(\frac{\partial u}{\partial x}\right)_{\text{right}} &\approx \frac{u_{n+1}-u_n}{d} \\
\left(\frac{\partial u}{\partial x}\right)_{\text{right}} &\approx \frac{u_{n+1}-u_n}{d} \\
F_{n,u} &\approx F_T\left[\left(\frac{\partial u}{\partial x}\right)_{\text{right}} - \left(\frac{\partial u}{\partial x}\right)_{\text{left}}\right] \\
\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) &\approx \frac{\left(\frac{\partial u}{\partial x}\right)_{\text{right}} - \left(\frac{\partial u}{\partial x}\right)_{\text{left}}}{d} = \frac{\left(\frac{F_n}{F_T}\right)}{d} \\
F_{n,u} &= F_T d\frac{\partial^2 u}{\partial x^2}
\end{aligned}
$$

Therefore the Momentum Principle gives us the following form of the wave equation in the case of transverse waves:

$$m_a\frac{\partial^2 u}{\partial t^2} = F_T d\frac{\partial^2 u}{\partial x^2}$$

If we divide by m_a, we obtain an equation that looks like the wave equation:

$$\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{F_T d}{m_a}\frac{\partial^2 u(x,t)}{\partial x^2}$$

Comparing with the general form of the wave equation, we see that

$$v^2 = \frac{F_T d}{m_a}$$

$$v = \pm\sqrt{\frac{F_T d}{m_a}}$$

In Section S3.9 is an optional derivation of the wave equation for light, which is a transverse wave.

A Macroscopic View of Transverse Waves

For concreteness we considered a single row of atoms, but inside a crystalline material (a material with an orderly lattice of atoms) there are of course many parallel rows of atoms, and there are spring-like forces between atoms in neighboring columns. For that reason transverse waves inside a block of metal are more complicated than waves along a hypothetical single row of atoms.

However, we can immediately apply our result to a macroscopic object. Consider a very low-mass string under tension F_T along which are attached beads of mass m_a, evenly spaced a distance d apart. The "linear density" (mass per unit length) is $\mu = m_a/d$. The speed of propagation of a transverse wave along the stretched string is

$$v = \sqrt{\frac{F_T d}{m_a}} = \sqrt{\frac{F_T}{\mu}}$$

This is valid as long as the wavelength is long compared to the bead spacing d. As the distance d between beads gets smaller and smaller, we have a nearly uniform-density string, so this is the speed of transverse waves in general along a string or rope.

It should be mentioned that in diagrams such as Figure S3.38 it is common to exaggerate greatly the vertical scale for u in order to be able to see the pulse clearly. Strictly speaking, the speed that we have calculated for transverse waves moving along a rope is valid only in the small-angle approximation.

> **QUESTION** If you double the tension in a rope, what happens to the speed of a transverse pulse traveling along the rope? If you double the tension in a metal bar, what happens to the speed of a longitudinal pulse (sound wave) traveling through the bar?

The speed of the pulse traveling along the rope increases by a factor of $\sqrt{2}$. This is quite different from the situation of a longitudinal sound wave in a bar of metal, where doubling the tension in the bar adds an amount s to the interatomic distance but doesn't change the equilibrium interatomic distance d, so the speed of sound doesn't change.

Sinusoidal Waves

A very important kind of wave is a sinusoidal wave. Many waves have a sinusoidal shape, and as we'll see in a later section on standing waves, it is possible to express mathematically other wave shapes as a sum of sinusoids.

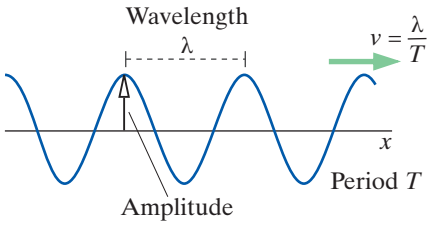


Figure S3.41 A sinusoidal wave.

Consider the following function shown in Figure S3.41, where A is called the amplitude, λ is the wavelength, and T is the period of the wave:

$$\begin{aligned} u(x,t) &= A\cos\left(\frac{2\pi x}{\lambda} - \frac{2\pi t}{T}\right) \\ &= A\cos\left(\frac{2\pi}{\lambda}\left(x - \frac{\lambda}{T}t\right)\right) \\ &= A\cos\left(\frac{2\pi}{\lambda}(x - vt)\right) \end{aligned}$$

This is a solution of the wave equation because it is a function of the quantity $(x - vt)$, with the speed of propagation $v = \lambda/T$ (in one period T the wave advances a distance of one wavelength λ). Moreover, when x advances by one wavelength λ or t advances by one period T, it is evident from the first form of the function shown above that the argument of the cosine advances by 2π rad (360°), and therefore the cosine function repeats.

The quantity $2\pi/\lambda$ is called the "wave number" k, and the quantity $2\pi/T$ is called the "angular frequency" ω. It is standard practice to write this solution of the wave equation in the form

$$u(x,t) = A\cos(kx - \omega t)$$

It is often useful to note the following:

$$\omega/k = (2\pi/T)/(2\pi/\lambda) = \lambda/T = v$$

Another quantity is often used, the frequency $f = 1/T$, which is the number of complete cycles in one second; the unit is hertz (Hz, cycles per second). Since $\omega = 2\pi/T$, $\omega = 2\pi f$.

> **QUESTION** What are the units of the wave number k? Of the angular frequency ω?

The wave number k has units of radians/m, or 1/m, and the angular frequency ω has units of radians/second, or 1/s. Note that radians are dimensionless.

Energy Propagation

A wave can carry energy from one place to a distant place without moving atoms from the initial location to the final location. In the case of sinusoidal waves we can calculate the rate at which energy is transferred, which is the power (energy delivered per second). Consider the kinetic energy and potential energy in a region of the wave that is one wavelength long. In one period T, all that energy is passed to the adjoining region or could be delivered to a device that is connected to the end of the waving object.

The kinetic energy of one atom in a row of atoms along which passes a sinusoidal wave is

$$\frac{1}{2}m_a\left(\frac{\partial u}{\partial t}\right)^2$$

The length of a row of atoms that we'll consider is λ, within which region there are $N = \lambda/d$ atoms since $Nd = \lambda$ is the total length. To add up the kinetic energies of all N atoms, K, we write the summation as an integral:

$$K = \int_0^\lambda \frac{1}{2}m_a\left(\frac{\partial u}{\partial t}\right)^2 \frac{dx}{d}$$

Dividing dx by d properly weights the various contributions, since the integral of dx/d is $\lambda/d = N$, the number of atoms whose kinetic energies we want to

add up. To put it another way, there is 1 atom per distance d, so dx/d is the number of atoms in a distance dx. For a sinusoidal wave $u = A\cos(kx - \omega t)$ we have

$$\frac{\partial u}{\partial t} = -\omega A \sin(kx - \omega t)$$

$$K = \int_0^\lambda \frac{1}{2} m_a (\omega^2 A^2 \sin^2(kx - \omega t)) \frac{dx}{d}$$

$$= \frac{1}{2} \frac{m_a \omega^2 A^2}{d} \int_0^\lambda \sin^2(kx - \omega t) dx$$

We can evaluate the integral by subtracting one trig identity from another:

$$\cos^2\theta + \sin^2\theta = 1$$

$$\cos^2\theta - \sin^2\theta = \cos(2\theta)$$

$$\sin^2\theta = \frac{1}{2}(1 - \cos(2\theta))$$

The integral over one wavelength of the cosine function is zero, leaving only the integral of dx:

$$\frac{1}{2} \int_0^\lambda (1 - \cos(2(kx - \omega t)) dx = \frac{\lambda}{2}$$

Another way to express this result is that the average value of $\sin^2$ over one period is 1/2, which is reasonable, since the square of the sine function varies between 0 and 1. The result for the kinetic energy in one wavelength of the wave is

$$K = \frac{1}{4} \frac{m_a \omega^2 A^2 \lambda}{d}$$

There is also potential energy in the wave associated with the stretch of each spring, with each spring contributing $(1/2)k_{s,\text{eff}}u^2$, where $k_{s,\text{eff}}$ is the effective spring stiffness associated with the spring forces exerted by neighboring atoms on an atom. For either longitudinal or transverse waves, the force on an atom can be expressed in terms of an effective spring stiffness for $u = A\cos(kx - \omega t)$:

$$F = m_a \frac{\partial^2 u}{\partial t^2}$$

$$= m_a(-\omega^2 u)$$

$$= -k_{s,\text{eff}} u$$

$$k_{s,\text{eff}} = m_a \omega^2$$

The length of atoms we'll consider is again λ, within which region there are $N = \lambda/d$ atoms since $Nd = \lambda$ is the total length. To add up the potential energies of all N atoms, U, we write the summation as an integral:

$$U = \int_0^\lambda \frac{1}{2} k_{s,\text{eff}} u^2 \frac{dx}{d}$$

Dividing dx by d properly weights the various contributions, since the integral of dx/d is $\lambda/d = N$, the number of atoms whose potential energies we want to add up. For a sinusoidal wave $u = A\cos(kx - \omega t)$ we have

$$U = \int_0^\lambda \frac{1}{2} m_a \omega^2 \sin^2(kx - \omega t)) \frac{dx}{d}$$

$$= \frac{1}{2} \frac{m_a \omega^2 A^2}{d} \int_0^\lambda \sin^2(kx - \omega t) dx$$

$$= \frac{1}{4} \frac{m_a \omega^2 A^2 \lambda}{d}$$

We see that the total potential energy in one wavelength is the same as the total kinetic energy, and their sum is this:

$$K + U = \frac{1}{2}\frac{m_a \omega^2 A^2 \lambda}{d}$$

Taking a macroscopic view, m_a/d is the linear mass density μ (kilograms per meter):

$$K + U = \frac{1}{2}\mu\omega^2 A^2 \lambda$$

The average power carried by the wave is the energy in one wavelength divided by the time for one wavelength to pass by, namely the period T:

$$\begin{aligned}\text{average power} &= \frac{1}{2}\frac{\mu\omega^2 A^2 \lambda}{T} \\ &= \frac{1}{2}\mu\omega^2 A^2 v\end{aligned}$$

The mass M of the waving object is the density ρ times the volume $L \times (\text{Area})$, where L is the length and Area is the cross-sectional area of the object. The mass M can also be written as μL, since μ is the mass per unit length. Therefore $M = \mu L = \rho L \times (\text{Area})$, and we can write the average power as

$$\frac{1}{2}\rho\,(\text{area})\,\omega^2 A^2 v$$

The average power per cross-sectional area (W/m^2) is

$$\frac{1}{2}\rho\omega^2 A^2 v$$

QUESTION Work out the units for this expression for average power. What should they be?

You should find that the units are joules/second, or watts. You can figure out the fundamental units for the joule by considering the units in the expression for kinetic energy, $(1/2)mv^2$.

QUESTION If you double the angular frequency, by what factor does the average power in the wave change? If instead you double the amplitude, how does the average power change?

In both cases the power will increase by a factor of four, because both ω and A are squared in the expression for average power.

In Chapter 23 we treated energy propagation in the case of light, which involves the Poynting vector.

Checkpoint 9 Diagrams of wave shapes often exaggerate the displacements by using an expanded vertical scale, to make it easier to see the wave shape. However, for the wave equation to correctly predict the speed of propagation of a transverse wave on a rope, it is important that the actual slope of the wave shape be small. What assumptions did we make in deriving the wave equation that required that the slope be small?

S3.5 STANDING WAVES

In this supplement and in Chapter 23 we have been studying traveling waves—waves whose crests travel through space at a speed that is characteristic of the type of wave (the speed of light for electromagnetic radiation, the speed of sound for sound waves, etc.). Next we will study waves that travel back and

forth within a confined space. Examples include elastic waves traveling back and forth along a length of string, or sound waves traveling back and forth inside an organ pipe. The behavior of such waves in these classical systems provides hints for understanding some quantum aspects of atoms.

Consider two sinusoidal traveling waves with exactly the same frequency that were initiated some time ago far to the left and far to the right and are heading toward each other (Figure S3.42). These waves might be electromagnetic waves in empty space, elastic waves propagating from two ends of a very long taut string, sound waves propagating from two ends of a long organ pipe, or water waves propagating from two ends of a lake. The description and analysis is practically the same in each of these very different physical situations. The wide applicability of the analysis we will develop is one of the reasons why we study waves. Recall that the the sum of solutions to the wave equation are also solutions (superposition property), so two waves approaching each other represent a valid solution of the wave equation.

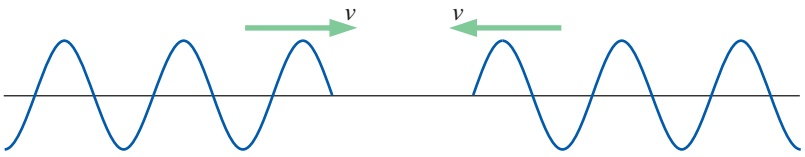


Figure S3.42 Colliding waves add up to create a standing wave.

When the waves start to overlap, they of course interfere with each other constructively or destructively, depending on the relative phases of the two waves in space. This interference leads to a remarkable effect. Look at the sequence of snapshots in Figure S3.43, taken one-eighth of a period apart. In the overlap region, the wave doesn't propagate! It just stays in one place, with the crests waving up and down but not moving to the left or to the right. This is something new, called a "standing wave."

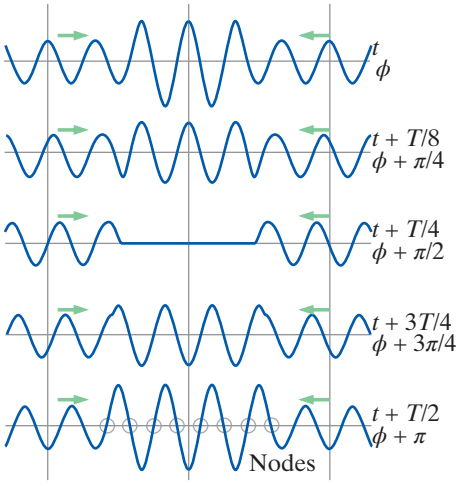


Figure S3.43 Two traveling waves of the same frequency produce a standing wave.

There are even instants when the wave is completely zero throughout the overlap region! At the instant that this occurs for a string, pieces of the string do have velocity (rate of change of displacement) and kinetic energy, so the waving does continue, even though there is momentarily no transverse displacement of the string anywhere in the region. In the case of electromagnetic radiation, the electric and magnetic fields are changing as they go through zero.

One of the important properties of a standing wave is the existence of what are called "nodes"—locations where the wave is zero *at all times*. Positions of nodes of the standing wave are circled on the bottom snapshot in Figure S3.42. Halfway between the nodes the standing wave periodically gets to full amplitude. These locations are called "antinodes."

Standing Waves vs. Traveling Waves

A single observer observing the wave as a function of time at one specific location that is not a node can't tell the difference between a traveling wave and a standing wave.

QUESTION Briefly explain why this is.

The single observer sees the wave going up and down at this location, which is just what it does when a traveling wave goes by. On the other hand, if the single observer is located at a node, the observer may conclude that nothing interesting is happening at all, or that this is a location where two traveling waves are destructively interfering with each other, as in two-slit interference, without forming a standing wave in space.

Only if we look at the complete pattern of waves in space and time can we see that a standing wave is really very different from a traveling wave. Yet there

are some properties that are the same for both kinds of wave. Look carefully at the time sequence of snapshots in Figure S3.42, and then answer the following important questions:

> **QUESTION** If the wavelength of the two traveling waves is λ, what is the wavelength (crest-to-crest distance) of the standing wave? If the frequency of the two traveling waves is f, what is the frequency (number of complete cycles per second) of the standing wave?

The wavelength and frequency of the standing wave are the same as for the traveling waves. This means that although the standing wave isn't going anywhere (left or right), we can still use the important relationship $v = f\lambda$ to relate frequency and wavelength for a standing wave through the speed v of a traveling wave.

Confined Waves: Standing Waves on a String

One of the ways to create a standing wave is to initiate a single traveling sinusoidal wave in a confined space, so that it continually reflects off the boundaries of this confined space, and the reflected wave interferes with the original wave. The frequency doesn't change in reflecting off a stationary boundary, so the original wave and its reflection are guaranteed to have the same frequency, which is necessary in order to be able to create a standing wave.

Examples of such confined waves include elastic waves on a string stretched between two supports, sound waves inside a pipe closed at both ends, and light waves reflecting back and forth between mirrors at each end of a laser. For concreteness we'll emphasize elastic waves on a taut string, but the analysis applies to many other quite different physical situations.

Figure S3.44 shows two possible standing waves on a taut string of length L that is tied to two fixed supports. The wavelength in the upper case is $\lambda = 2L$, as can be verified by noting that the length L contains half a wavelength. The wavelength in the lower case is $\lambda = L$.

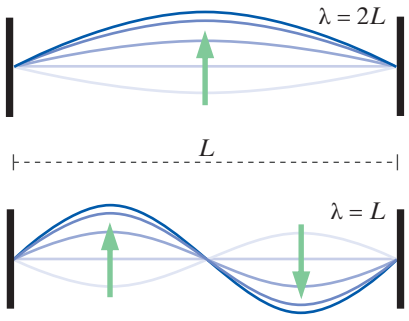


Figure S3.44 Two of the possible standing waves on a taut string of length L.

> **QUESTION** Why can't you have a standing wave on this string with wavelength $3L$, or $1.1L$, or $0.95L$?

These wavelengths don't fit because the waves aren't zero at the ends of the string, where the string is tied to fixed supports. You have discovered that the possible standing waves on a string are "quantized." Not just any old wavelength will do. The possible wavelengths of a standing wave on a taut string are completely determined by the distance between the supports. The same important principle applies to possible wavelengths of standing sound waves in an organ pipe or possible wavelengths of standing light waves in a laser.

STANDING WAVES ARE QUANTIZED

Only certain wavelengths (and frequencies) are possible.

Enumerating the Possible Standing Waves

Let's generate a more complete catalog of the possible wavelengths for standing waves on a taut string.

> **QUESTION** In terms of the length L, write down the six longest possible wavelengths. You already know the first two: $\lambda_1 = 2L$ and $\lambda_2 = L$.

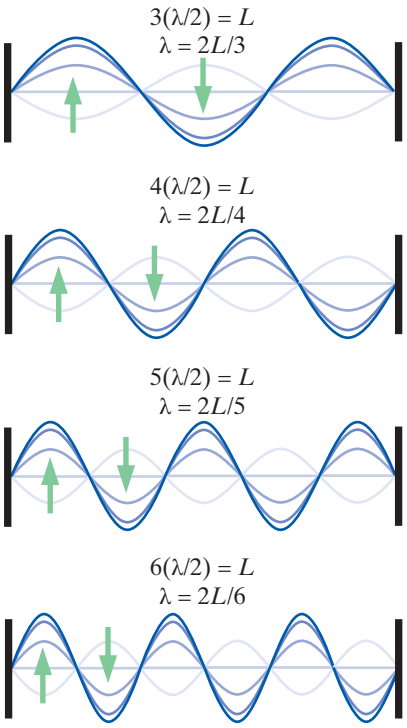


Figure S3.45 The next four possible standing waves.

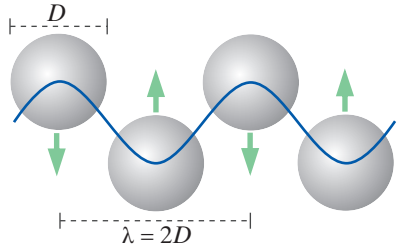


Figure S3.46 The shortest possible wavelength (highest possible frequency).

Evidently we have $\lambda_3 = 2L/3$, $\lambda_4 = 2L/4$, $\lambda_5 = 2L/5$, and $\lambda_6 = 2L/6$. The waves are shown in Figure S3.45. A quick way to determine the possible wavelengths is to note that in the length L there must be an integer number n of half-wavelengths, so $L = n(\lambda_n/2)$, which means that $\lambda_n = 2L/n$. As long as the amplitude isn't too large, or the wavelength too small, the speed of propagation v of traveling waves on a taut string turns out to be independent of the wavelength.

> **QUESTION** In terms of this speed v, what are the frequencies of the six lowest possible standing-wave frequencies?

Solving for $f_n = v/\lambda_n = nv/(2L)$, we have $f_1 = v/(2L)$, $f_2 = 2v/(2L)$, $f_3 = 3v/(2L)$, $f_4 = 4v/(2L)$, $f_5 = 5v/(2L)$, and $f_6 = 6v/(2L)$. These different standing waves are called "modes." The technical use of the term is that a mode of oscillation of a system is one characterized by a standing wave that is sinusoidal in time with the same frequency at every location, and that necessarily satisfies the boundary conditions (in our case, that the wave is zero at the ends of the string). A mode may also be sinusoidal in space at a particular instant of time, as is true for a string, but it need not be spatially sinusoidal in more complicated systems, such as the two-dimensional surface of a drum.

We've seen that there is a longest possible wavelength $2L$ and a lowest possible frequency $v/(2L)$. Is there a shortest possible wavelength, corresponding to a highest possible frequency? Yes. If the distance between atoms in the string is D, the shortest theoretically possible wavelength for a standing elastic wave is $2D$, with adjacent atoms oscillating out of phase with each other (Figure S3.46). Any shorter wavelength has no physical significance.

As is the case for many other systems, a vibrating string has a number of modes (number of different possible standing-wave patterns) that is not infinite but finite.

Building Up to the Steady State

You may have had the experience of building up a standing wave in a string or rope. You tie one end of the rope to something, pull it taut, then shake your end of the rope up and down. If you shake at one of the mode frequencies, a large oscillation quickly builds up, because the traveling wave you send down the rope reflects off the other end in such a way as to produce a standing wave through constructive interference. However, if you try to shake the rope at a frequency that is not right for any mode, you don't build up a large oscillation, because you don't get constructive interference.

As you continue to put energy into the rope by shaking it at an appropriate frequency, the standing wave grows and grows. What limits its growth? What determines the final amplitude of the standing wave?

At first, almost all of your energy input goes into making the standing wave grow. As the amplitude increases, however, so does the rate of energy losses due to air resistance and "internal friction" in the rope itself associated with bending the material. The air resistance grows because air resistance is larger at higher speeds, and the rope has higher transverse speeds when the amplitude of the standing wave is larger. Internal friction inside the rope also grows as the bending angle of the rope gets larger with larger amplitude.

At some particular amplitude the losses to air resistance and internal friction in each cycle have grown to be exactly equal to the energy input you make each cycle. During each cycle you simply make up for the losses, but no further growth of the wave occurs. You have reached a steady state.

If the losses are relatively small, you don't have to do much to keep the wave going. In fact, you may have experienced keeping a rope oscillating with a sizable amplitude while your hand hardly moves. In a low-loss situation,

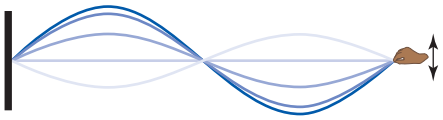


Figure S3.47 With low losses, very little motion of the hand is sufficient to maintain the steady state.

your hand is very nearly a node, in that your hand moves a very tiny amount compared with the large motion at one of the antinodes (Figure S3.47).

Checkpoint 10 Given the loss mechanisms we have outlined, which should be easier to build up and sustain: a long-wavelength (low-frequency) mode or a short-wavelength (high-frequency) mode? If you have the opportunity, try this yourself with a rope.

Superposition of Modes

We've established that standing waves (modes) are quantized for oscillating strings. Is that all a string can do—be in one of these modes? No! According to the superposition principle, it ought to be possible for a string to be oscillating in a manner described by the superposition of two or more modes. What would that look like? Would it be a standing wave?

Figure S3.48 shows you a time sequence of the superposition of the two lowest-frequency modes (wavelengths $2L$ and L). Both modes have been given the same amplitude. The individual modes are shown with light lines, and their sum is shown with a colored line.

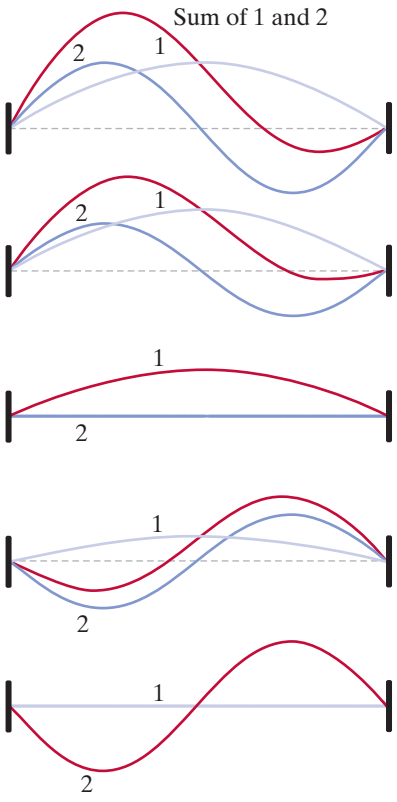


Figure S3.48 A confined wave that is the superposition of the two lowest-frequency modes.

QUESTION Is this motion a standing wave? A traveling wave? Why or why not?

Neither. It is not a standing wave, nor is it a simple traveling wave, because crests move both to the left and to the right. Also, there isn't a "wavelength," which is something that both standing and traveling waves have. Despite the complexity of this motion, it is at least periodic—the motion repeats after a time. To understand how the motion repeats, in Figure S3.48 it is instructive to follow mode 1 from top to bottom, then follow mode 2 from top to bottom.

QUESTION If the period of the longest-wavelength mode is $T(=2L/v)$, what is the period of this two-mode oscillation?

After T, the low-frequency mode has gone through one complete cycle, and the second mode has gone through two complete cycles, and we're back at the original configuration. Therefore the period is T. Another way to see this is that in Figure S3.48, mode 1 goes through a quarter cycle (from maximum positive to zero), while mode 2 goes through two quarter cycles (from maximum positive on the left to maximum negative on the left). Extending the set of diagrams to four times as much time would bring both modes back to the starting configuration.

Combinations of large numbers of modes can represent pretty complicated motions. Figure S3.49 is a time sequence of the superposition of particular amplitudes of ten different modes, chosen in such a way as to approximate a "square wave" initially.

How complicated can this get? It can be mathematically proven that any function $f(x)$ that is zero at the ends of the string, $f(0) = 0$ and $f(L) = 0$, can be created by adding up carefully chosen amounts of a (possibly infinite) number of modes. In a section on Fourier analysis at the end of this supplement we show you how we determined the amplitudes of each of the ten modes we used to approximate the wave shown in Figure S3.49.

This provides an extremely powerful way of analyzing the behavior of confined waves. First figure out what the modes are (the possible sinusoidal standing waves). Then figure out what combination of modes can make a particular shape $f(x)$ at time $t = 0$. What will happen from then on can be calculated simply by following the sinusoidal behavior of the individual modes, each with their own characteristic frequencies, and adding up the individual contributions of the various modes.

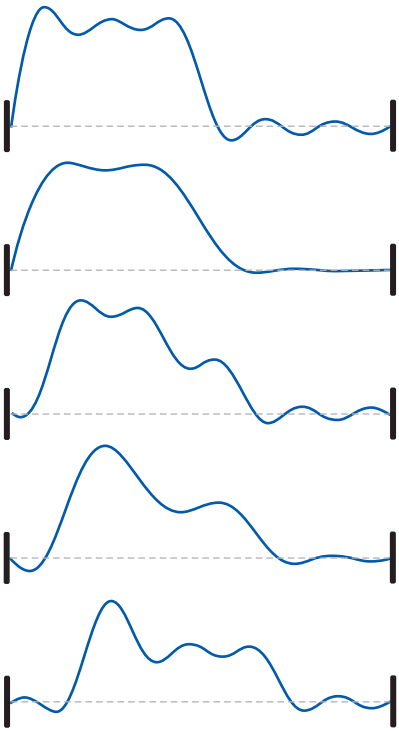


Figure S3.49 An approximation to an initial square wave, using ten modes.

Musical Tones

When a violin string is bowed, it oscillates with a superposition of various modes, and these oscillations drive the body of the violin to oscillate, which alternately compresses and rarefies the air to produce the sound waves that you hear. Usually the largest-amplitude mode is the longest-wavelength (lowest-frequency) mode, which corresponds to what is called the "pitch" or "fundamental frequency." A pure single-frequency tone sounds very boring. The rich musical tone you hear from a violin is due not just to the fundamental frequency but to the superposition of many higher-frequency modes that are also present. Mode frequencies that are simple integer multiples of the fundamental frequency are called "harmonics." It is a peculiarity of the human ear and brain that we find the superposition of harmonics very pleasing. The superposition of frequencies that are not simple integer multiples of a fundamental frequency sounds much less pleasing.

Nonharmonic Oscillations

The example of a string is highly representative of a very broad range of standing-wave phenomena. However, we should alert you to some issues that are not illustrated by the relatively simple behavior of a string. First, while the geometry of the system determines the possible wavelengths (in order to satisfy the boundary conditions), the different wavelengths need not be simply $1/n$ times the longest wavelength. A good example is a drum. The wavelengths of the two-dimensional waves on a drumhead are determined by the requirement that the wave be zero at the drumhead boundary. The relationship among the various mode wavelengths is rather complicated, not simply $1/n$ times the longest wavelength. The fact that the wavelengths of drum modes are not simply related to each other (are not harmonics) corresponds to the fact that a drum does not produce a musical tone such as is produced by a violin.

Speed Dependent on Frequency

Second, the speed of propagation in many systems is not constant but depends on the frequency (or wavelength). Since $f = v/\lambda$, in such systems the frequency of a mode need not be simply inversely proportional to the wavelength of that mode, as it is for a string. Such systems are called "dispersive" because traveling waves of different frequencies spread out (disperse) from each other due to differences in their speeds. The analysis in terms of modes for dispersive systems is still valid. It's just that when you calculate $f = v/\lambda$ for a particular λ, you have to consider the fact that v isn't a constant.

Nonlinear Systems

Third, we relied on the superposition principle to combine modes simply by adding them up. More fundamentally, we implicitly assumed that if we double the amplitude of a mode, the frequency won't change. Such systems are called "linear" systems (because the amplitude of the superposition of two waves of the same mode is just the same mode, with an amplitude that is the sum of the amplitudes of the two waves). However, many real systems are only approximately linear in this sense. In "nonlinear" systems, the superposition of two waves of the same frequency may lead to a wave with a different frequency.

This can even happen with a string. We have seen that the speed of propagation v for a string is proportional to the square root of the tension in the string. For small amplitudes, the string is not stretched very much more than it was at rest, and the tension is not changed very much. But for large amplitudes, the string may be stretched quite a bit more than it is when at rest, and that means a change in v, which means a change in frequency for the same

wavelength. Hence two medium-sized waves might add up to a large wave with a different frequency.

Confined Waves and Quantum Mechanics

We have been considering quantization in classical (non-quantum) systems. We found that standing waves in confined regions couldn't have just any old wavelength, because the possible wavelengths are quantized by the constraints of the boundaries.

There is an analogous situation in the quantum world of atoms, due to the fact that particles such as electrons can have wavelike properties. For example, we found that electron diffraction demonstrates the startling fact that electrons have wave properties as well as particle properties. This implies that there could be three-dimensional standing waves of an electron in a hydrogen atom, because the electron is confined to a region near the proton, due to the electric attraction. A hydrogen atom does indeed have quantized states analogous to standing-wave modes. Each of these states has a specific energy—the energy is quantized. A hydrogen atom can drop from a higher-energy to a lower-energy state with the emission of energy in the form of light, and a study of the light emitted by atoms lets us determine the quantized energies of atomic states.

S3.6 WAVE AND PARTICLE MODELS OF LIGHT

We have established that a variety of phenomena can be explained by the wave properties of electromagnetic radiation. Yet in Chapters 8 and 10 we explained a variety of phenomena by a particle (photon) model of light. As a first step in comparing the wave and particle models of light, let's consider the predictions of the two models of light in two different experimental situations: the photoelectric effect and the collision of a beam of light with free electrons (Compton scattering).

The Photoelectric Effect

The mobile electrons in a metal, although quite free to move around inside the metal, are nevertheless bound to the metal as a whole: they don't drip out! It takes some work to move an electron from just inside to just outside the surface of the metal. At that point (electron just outside the surface), the metal would now be positively charged and would attract the electron, so you would need to do additional work to move the electron far away. The total amount of work required to remove an electron from a metal is often called the "work function" but could just as well have been called the "binding energy." We will refer to it with the symbol W. Typical values for common metals are a few electron volts (eV).

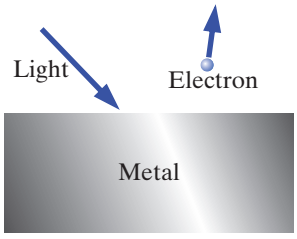


Figure S3.50 The photoelectric effect: light can eject an electron from a metal.

Prediction of the Wave Model

One way to supply enough energy to remove an electron from a metal is to shine light on the metal (Figure S3.50).

> **QUESTION** According to the wave model of light, how could a beam of light have sufficient energy W to overcome the binding energy and remove an electron from a piece of metal?

Since the energy per square meter (intensity) of an electromagnetic wave is proportional to E^2, we would need a high-intensity electromagnetic wave, in which the amplitude of the electric field is very large, to supply enough energy. Alternatively, if the energy from the electromagnetic wave were not quickly

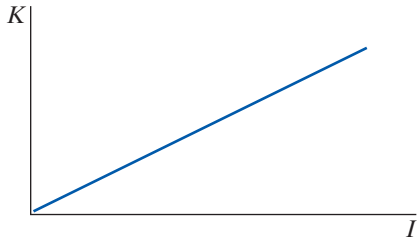


Figure S3.51 The wave model of light predicts that the kinetic energy of an ejected electron should be proportional to the intensity of light hitting the metal.

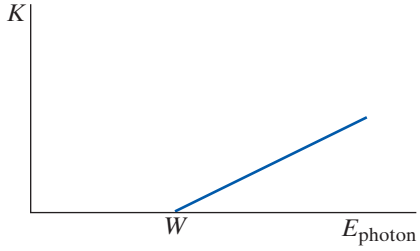


Figure S3.52 The particle model predicts that above a certain threshold energy W, the kinetic energy of an ejected electron should be proportional to $(E_{photon} - W)$.

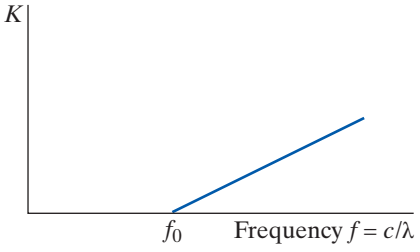


Figure S3.53 What is actually observed is that above a minimum frequency f_0, the kinetic energy of ejected electrons is proportional to $(f - f_0)$.

dissipated into thermal energy, we might be able to shine a low-intensity beam on the metal for a long time to supply enough energy.

> **QUESTION** According to the wave model of light, would we need to use electromagnetic radiation of any particular wavelength, or would any wavelength be adequate?

According to the wave model, the energy density of an electromagnetic wave depends on E^2, the square of its amplitude, not on its wavelength. All we need is an intense beam of light—it could have any wavelength (or color). We might therefore expect a plot of the kinetic energy of an ejected electron vs. intensity of light hitting the metal to be linear, as in Figure S3.51.

Prediction of the Particle Model

Recall from the previous volume that according to the particle model a beam of light is a collection of particles traveling at speed c, each with zero rest mass but with a fixed energy and momentum.

> **QUESTION** What does the particle model of light predict would be required to eject an electron from a metal?

The particle model suggests that the energy of an individual photon is what is important, since the absorption of a single photon of the appropriate energy would be adequate to free an electron from the metal (and give it some kinetic energy). This model predicts that an extremely intense beam (many photons per square meter) will have no effect if the energy of the individual particles is too low (Figure S3.52).

What Is Observed?

In fact, what is observed experimentally is that the wavelength of the light that falls on the metal surface determines whether or not an electron will be liberated from the metal (Figure S3.53). If the wavelength of the light is too large (the frequency is too low), nothing happens, even if the beam is very intense. However, as the wavelength of the radiation is decreased (frequency increased), at one particular wavelength we begin to see electrons liberated from the metal. As the wavelength of the incident light is decreased even further, the kinetic energy of the electrons liberated from the metal increases.

What happens if we use a very low intensity beam? At low intensity, long-wavelength light does not eject electrons, no matter how long we wait. However, as predicted by the particle model, with a low intensity beam of short wavelength light, we sometimes see an electron ejected almost immediately.

For light of short enough wavelength, we do however observe that the rate at which electrons are ejected from the metal is proportional to the intensity of the light.

Unifying the Two Models

How do we reconcile our models with the experimental observations? First, it seems clear that this phenomenon is essentially a quantum one. A certain amount of energy must be delivered in one chunk, or packet, to the system, in order to raise the system to an energy state in which the electron is unbound. Thus, the particle model of light seems appropriate here.

The surprising thing, however, is that the wavelength of the incident light is important. Wavelength is certainly a property of a wave—how can this be reconciled with a particle model of light?

Evidently, we need to model light as something with both wave-like and particle-like properties. The wave nature of light explains the interference

Figure S3.54 A photon can be described as a particle with momentum and energy, yet strangely enough with a wavelength as well.

phenomena we have observed. The particle nature of light explains how a fixed amount of energy can be delivered to a system by a photon. It turns out to be possible to describe photons as particles that have momentum and energy, yet strangely enough with a wavelength as well (Figure S3.54).

Energy and Wavelength

From experimental observations of the photoelectric effect, it is found that the energy of a photon is proportional to the frequency and inversely proportional to the wavelength. Quantitatively, the relation is

$$E = hf = \frac{hc}{\lambda}$$

where $h = 6.6 \times 10^{-34}$ J-s is Planck's constant, which you may remember from our study of quantized harmonic oscillators in Volume I.

Intensity and Number of Photons/Second

Since the number of electrons per second ejected is proportional to the intensity of light of sufficiently small wavelength, apparently in the particle model the intensity of light is proportional to the number of photons per second striking a surface.

Applying the Model

> **QUESTION** If the work function (binding energy) $W = 3$ eV for a certain metal, what is the minimum energy a photon must have to eject an electron?

Evidently the photon must have an energy of 3 eV or more. If the photon has less energy, it can't knock an electron out of this metal.

> **QUESTION** What is the wavelength of a photon with this much energy?

$$\lambda = \frac{hc}{E} = \frac{(6.6 \times 10^{-34}\ \text{J·s})(3 \times 10^{8}\ \text{m/s})}{(3\ \text{eV})(1.6 \times 10^{-19}\ \text{J/eV})} = 4.12 \times 10^{-7}\ \text{m} = 412\ \text{nm}$$

This is near the violet end of the visible spectrum, which extends from about 400 nm (3.1 eV) to about 700 nm (1.8 eV). For a metal with $W = 3$ eV, light in most of the visible spectrum or in the infrared cannot eject an electron. UV (ultraviolet) light, with wavelength shorter than 400 nm (photon energy greater than 3.1 eV), can knock electrons out of this metal.

> **QUESTION** What if you expose this metal to light whose photon energy is 4 eV (in the UV)? How much kinetic energy would you expect the ejected electron to have when it was far from the metal?

Since it takes 3 eV just to remove the electron, a 4 eV photon can eject an electron with 1 eV to spare, so the electron will have a maximum kinetic energy of 1 eV (it could have less if it loses some energy on its way out of the solid).

Experiments have verified this simple particle model for the photoelectric effect. The main points are these:

- There is a minimum photon energy W required to eject an electron from a metal; photons of lower energy (longer wavelength) do not eject electrons.
- Photons with excess energy eject electrons whose maximum kinetic energy is equal to the photon energy minus W.

Historically, at the beginning of the 20th century the wave model of light was the accepted model. It was Einstein who first proposed the particle interpretation of the photoelectric effect, in 1905. (In the same year Einstein also published the theory of special relativity, and he also correctly analyzed Brownian motion, the random motion of small objects due to molecular collisions, which led to the first accurate determination of atomic sizes. A truly remarkable year!)

Checkpoint 11 If a particular metal surface is hit by a photon of energy 4.3 eV, an electron is ejected, and the maximum kinetic energy of the electron is 0.9 eV. What is the work function W (binding energy) of this metal?

Compton Scattering

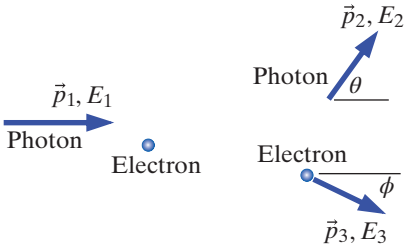


Figure S3.55 Compton scattering: a photon scatters off a stationary electron.

Figure S3.55 is a diagram of a collision between a photon and a stationary electron. This process is called "Compton scattering" in honor of the physicist who showed experimentally that this process has both particle and wave aspects. If the electron recoils with some kinetic energy, the photon necessarily has lost energy. In the wave model of the interaction, loss of energy corresponds to decreased amplitude but the same frequency, whereas in the wave–particle model a loss of photon energy corresponds to a longer wavelength. Compton showed experimentally that the combined wave–particle model did correctly predict the change of wavelength as a function of the scattering angle of the electron. This was one of the most compelling early pieces of evidence for the combined wave–particle nature of light.

In Volume 1 you used the Momentum Principle and the Energy Principle to study particle collisions, and you treated photons as particles with zero rest mass. The general relationship between energy and momentum for a particle is $E^2 = (pc)^2 + (mc^2)^2$. For zero-mass particles such as photons (and probably neutrinos), this reduces to $E = pc$. We will use this in an analysis of Compton scattering.

Let's use the combined wave–particle model of light to predict the change in wavelength $\Delta\lambda$ for a photon that collides with a free electron. As indicated in Figure S3.55, let the energy of the incoming photon be E_1, the outgoing photon energy be E_2, and the energy of the recoil electron be E_3, where $E_3 = mc^2 + K$, K is the kinetic energy of the recoil electron, and m is the mass of the electron.

$E_1 + mc^2 = E_2 + E_3$	Energy Principle
$E_3 = E_1 - E_2 + mc^2$	Solve for recoil electron energy
$\vec{p}_1 = \vec{p}_2 + \vec{p}_3$	Momentum Principle
$\vec{p}_3 = \vec{p}_1 - \vec{p}_2$	Solve for recoil electron momentum

To get the square of the magnitude of a vector, we take the dot product of the vector with itself:

$\vec{p}_3 \bullet \vec{p}_3 = (\vec{p}_1 - \vec{p}_2) \bullet (\vec{p}_1 - \vec{p}_2)$	Dot product
$p_3^2 = p_1^2 + p_2^2 - 2\vec{p}_1 \bullet \vec{p}_2 = p_1^2 + p_2^2 - 2p_1 p_2 \cos\theta$	Expand
$(p_3c)^2 = (p_1c)^2 + (p_2c)^2 - 2(p_1c)(p_2c)\cos\theta$	Multiply by c^2
$E_3^2 - (mc^2)^2 = E_1^2 + E_2^2 - 2E_1E_2\cos\theta$	Substitute $(pc)^2 = E^2 - (mc^2)^2$
$(E_1 - E_2 + mc^2)^2 - (mc^2)^2 = E_1^2 + E_2^2 - 2E_1E_2\cos\theta$	Substitute E_3

After expanding the squared parenthesis and simplifying, we obtain this:

$$mc^2(E_1 - E_2) = E_1E_2(1 - \cos\theta)$$

$$mc^2\left(\frac{1}{E_2} - \frac{1}{E_1}\right) = (1 - \cos\theta) \qquad \text{Divide by } E_1E_2$$

$$mc^2\left(\frac{\lambda_2}{hc} - \frac{\lambda_1}{hc}\right) = (1 - \cos\theta) \qquad \text{Use wave–particle relation } E = \frac{hc}{\lambda}$$

$$\lambda_2 - \lambda_1 = \frac{h}{mc}(1 - \cos\theta) \qquad \text{Compton scattering prediction}$$

Here we have a quantitative prediction for the increase in wavelength of the scattered photon, as a function of the scattering angle θ. The quantity

$$\frac{h}{mc} = \frac{(6.6 \times 10^{-34}\ \text{J·s})}{(9 \times 10^{-31}\ \text{kg})(3 \times 10^{8}\ \text{m/s})} = 2.4 \times 10^{-12}\ \text{m}$$

is called the Compton wavelength. To make it feasible to measure the wavelength increase, one must use light whose wavelength is comparable to the Compton wavelength. Such light is in the x-ray region of the electromagnetic spectrum.

To study this kind of collision, it is necessary to have a beam of monochromatic (single wavelength) x-rays. However, most x-ray beams start out with a range of wavelengths.

> **QUESTION** Given what you know about x-ray diffraction, can you think of a way to produce a single-wavelength beam?

If a beam of x-rays interacts with a crystal, re-radiated x-rays of a given wavelength will be particularly intense at an angle satisfying the x-ray diffraction condition. Such a single-wavelength beam could be used in the experiment.

> **QUESTION** How might one measure the wavelength of the x-rays after they interact with an electron?

Again, they can be measured by using x-ray diffraction from a single crystal. If we know the spacing of atoms in a particular crystal, measuring the angles of diffraction maxima will tell us the wavelength of the x-rays.

Compton did the experiment, shooting x-rays of known, single wavelength at a block of metal (which contains lots of nearly free electrons) and measuring the wavelength of the scattered x-rays at various scattering angles. He found that the wavelength did indeed increase by the amount predicted by the analysis given above.

To summarize, a particle view of light was used to predict the change in energy of the photon, and the wave–particle relation $E = hc/\lambda$ was used to convert the energy change to a change in wavelength. The experiment showed that the combination of the particle model and the wave model did indeed predict what actually happens in this process. Neither a pure wave model of the process nor a pure particle model would make the correct prediction.

Checkpoint 12 In a Compton scattering experiment, if the incoming x-rays have wavelength 5×10^{-11} m, what is the wavelength for x-rays observed to scatter through 90°?

Two-Slit Interference Revisited

We saw previously that if a beam of monochromatic light passes through two slits, an interference pattern is formed on a distant surface. Using the wave

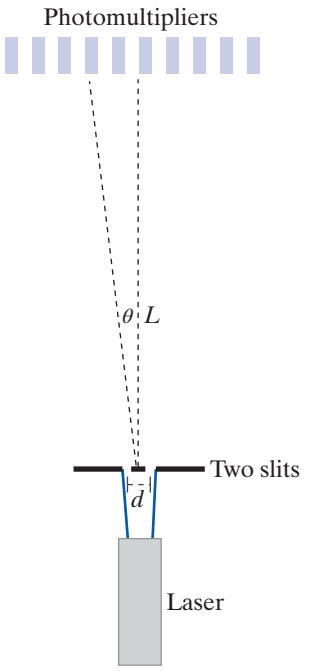


Figure S3.56 A two-slit experiment with a very low intensity beam, using photomultipliers as photon detectors.

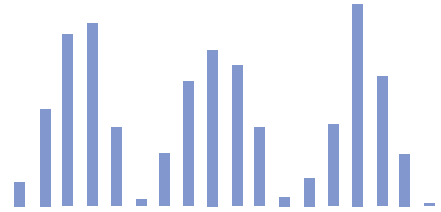


Figure S3.57 Number of photons detected vs. angle for a low-intensity beam passing through two slits.

model of light, and applying the superposition principle, we were able to predict accurately the angles at which maxima and minima would occur in the interference pattern.

Suppose we shine a beam light through two slits, as before, but this time we decrease the intensity of the light until the beam is extremely weak. The amount of light hitting a distant screen is too small to be detected by human eyes, so we use a different kind of detector: a photomultiplier (Figure S3.56). The functioning of a photomultiplier tube is based on the photoelectric effect. If a single photon ejects an electron from a metal surface, the resulting one-electron "current" is amplified electronically into a signal strong enough to detect. Photomultiplier tubes can detect individual photons. We will put an array of photomultipliers at the location of the screen where we previously saw an interference pattern with a strong beam of light, and make a plot of number of photons detected vs. location.

> **QUESTION** What would you expect the appearance of the resulting plot to be?

Although individual photons are detected at particular locations, over time an interference pattern builds up, like the one observed with an intense beam (Figure S3.57). This pattern, however, builds up statistically, with photons frequently detected at locations of maxima and never detected at locations of minima. Despite the fact that we are detecting individual particles, we still observe a wavelike phenomenon.

It is clear from this and other experiments that a complete model of light must have both a wave-like and a particle-like character, as we concluded earlier.

Electron Diffraction

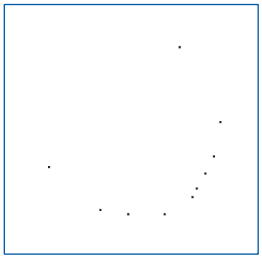
If we shoot a beam of x-rays at a single atom, re-radiation of x-rays by accelerated electrons in the atom goes in nearly all directions. However, if we shoot a beam of x-rays at a crystal made up a huge number of atoms arranged in a three-dimensional array, there is constructive interference of the re-radiation only in certain special directions for which the condition $2d\sin\theta = n\lambda$ is true. This is very much a wave phenomenon.

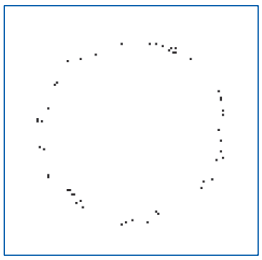
Earlier in this supplement we discussed x-ray diffraction by polycrystalline powders, in which there is strong constructive interference for those tiny crystals that happen to be oriented at the appropriate angle for the condition $2d\sin\theta = n\lambda$ to be true. The result is rings of re-radiated x-rays surrounding the incoming x-ray beam. The larger the wavelength λ, the larger the angle θ, and the larger the rings seen on the x-ray film.

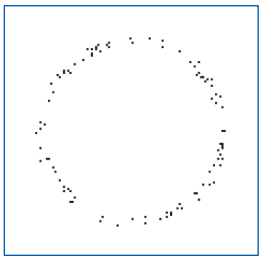
We would expect very different behavior if we used particles instead of waves. Suppose that instead of using x-rays we shoot electrons at a single atom. There is a complicated electric interaction between an incoming electron and the atom, including polarization of the atom by the electron. The electron is deflected by the collision with the atom and can go in almost any direction, depending on how close to center it hits the atom.

> **QUESTION** If electrons were shot at a crystal made up of a huge number of atoms arranged in a three-dimensional array, what would you expect to see?

We certainly don't expect to see any interference effects, since we assume that an electron is a particle, not a wave. Nevertheless, if we shoot a beam of electrons into a polycrystalline powder, we get rings of electrons, just as with x-ray diffraction, with the angles of the rings described by the usual x-ray diffraction equation! Experimentally, it appears that electrons also can act like waves.


After 10 electrons


After 50 electrons


After 100 electrons

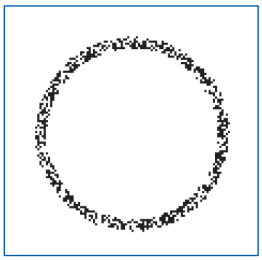

After 1000 electrons

Figure S3.58 Longer and longer exposures to electrons passing through a polycrystalline powder.

Moreover, an electron not only exhibits wave-like interference phenomena but even appears to interfere with itself! If you send in electrons one at a time (a very low-current beam of electrons), and detect the electrons one at a time with an array of detectors (or a photographic film), you get an interference pattern, but built up in a statistical way. Figure S3.58 shows what you observe with longer and longer exposures of film with a low-current beam of electrons hitting a polycrystalline powder. Each dot represents a location where an electron hit the film.

This is intriguing! You detect individual electrons, like particles. Yet if you wait long enough, you find that the pattern of detection looks like what you get with x-rays, which are waves. We say that electrons have properties that are both particle-like and wave-like.

In these electron experiments the condition for (statistical) constructive interference is $2d\sin\theta = n\lambda$, just the same as for x-ray diffraction. What is the wavelength λ of the electrons? It was predicted theoretically by de Broglie and verified experimentally by Davisson and Germer that

$$\lambda = \frac{h}{p}$$

where p is the momentum of the electron, and h is Planck's constant, an extremely small quantity measured to be $h = 6.6 \times 10^{-34}$ J · s. Note that since for a photon $E = pc$, this relation is valid for photons too:

$$\lambda = \frac{hc}{E} = \frac{h}{p}$$

This relationship is easily demonstrated with an electron-diffraction apparatus in which you can control the momentum of the incoming electrons by varying the accelerating voltage ΔV. We calculate the momentum corresponding to a given kinetic energy K:

$$K = \frac{p^2}{2m}$$

so $p = \sqrt{2mK}$, where $K = e\Delta V$ if the electrons start from rest. Therefore we predict this wavelength for the electron:

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2me\Delta V}}$$

QUESTION Based on this relation and on the diffraction relation $n\lambda = 2d\sin\theta$, how would you expect the spacing of electron diffraction rings to change if you increase the accelerating voltage?

Experimentally we find that if the accelerating voltage ΔV is quadrupled, the electron wavelength is decreased by a factor of 2, and the diffraction rings shrink due to $\sin\theta$ being reduced by a factor of 2.

Electron diffraction is used extensively to study the structure of materials. It has an advantage compared with x-ray diffraction in that electron beams can be focused easily to a very small spot using electric and magnetic fields, whereas this is difficult to do with x-rays. For this reason, electron diffraction is particularly useful in studying small regions of a sample of material. In actual practice, rather high-energy electrons are used in order to achieve good penetration of the material, with typical accelerating potentials of several hundred thousand volts. That means that the wavelength λ is much smaller than a typical interatomic spacing d. Consequently θ is quite small, and the electrons are sent in nearly parallel to the atomic planes of interest.

Neutron Diffraction

Perhaps electrons and photons are very special in having both particle and wave aspects? Not really. Similar diffraction patterns have been observed when entire helium atoms are shot at a crystal, and the wavelength of the helium atom is again $\lambda = h/p$ (but with a much larger mass m).

A particularly important case is neutron diffraction. Neutrons have no charge and only a very small magnetic moment, so neutrons have no electric interactions with electrons and a small magnetic interaction. However, neutrons do have a strong nuclear interaction with protons and other neutrons, as is evident by the presence of neutrons in the nuclei of atoms. If you shoot a single neutron at a single atom, the neutron may be hardly affected by the electrons but can be deflected by the nucleus through just about any angle. When you shoot neutrons at a crystal, the usual equation $2d\sin\theta = n\lambda$ for constructive interference applies.

What Is It That Is "Waving"?

Electromagnetic radiation consists of waves of electric field. Interference of electromagnetic radiation is due to superposition of electric fields, and the intensity is proportional to the square of the field amplitude.

In the case of electron or neutron diffraction, what is it that is "waving"? It turns out that particles can be described by abstract "wave functions" whose amplitude squared at a location in space is the probability of finding the particle there. The same equation $2d\sin\theta = n\lambda$ that predicts the locations of rings of intensity for x-ray diffraction also predicts the locations of rings of probability for electron or neutron diffraction. Classical interference of light is therefore a good foundation for understanding quantum interference of particles—but the interpretation changes from classical intensity to quantum probability.

Checkpoint 13 Roughly, what is the minimum accelerating potential ΔV needed in order that electrons exhibit diffraction effects in a crystal?

S3.7 *FOURIER ANALYSIS

How did we figure out what amount of each of the ten modes to superimpose to get the complicated behavior shown in Section S3.5? In principle we could have done it by trial and error—just try various combinations of modes until we get the shape we want. However, you can easily imagine that it would be very tedious to try to get this right merely by trial and error.

We'll outline the general scheme for doing this, then apply it to our specific case. We want to express the initial shape of the string, $f(x)$, as a sum of appropriate amounts of many modes, which are sine functions with wavelengths $2L$, $2L/2$, $2L/3$, and so on:

$$f(x) = A_1 \sin\left(2\pi\frac{x}{2L}\right) + A_2 \sin\left(2\pi\frac{2x}{2L}\right) + A_3 \sin\left(2\pi\frac{3x}{2L}\right) + \cdots$$

We need to adjust the coefficients A_1, A_2, A_3, and so on in order to choose the appropriate superposition of modes that is equal to $f(x)$. In order to figure out what the nth coefficient should be, consider the following integral:

$$\int_0^L f(x)\sin\left(2\pi\frac{nx}{2L}\right)dx =$$
$$\int_0^L \left[A_1 \sin\left(2\pi\frac{x}{2L}\right) + A_2 \sin\left(2\pi\frac{2x}{2L}\right) + \cdots\right] \sin\left(2\pi\frac{nx}{2L}\right)dx$$

It can be shown by using trig identities that the following is true:

$$\int_0^L \sin\left(2\pi\frac{kx}{2L}\right)\sin\left(2\pi\frac{nx}{2L}\right)dx = \frac{L}{2} \quad \text{if } k = n$$
$$\int_0^L \sin\left(2\pi\frac{kx}{2L}\right)\sin\left(2\pi\frac{nx}{2L}\right)dx = 0 \quad \text{if } k \neq n$$

Therefore the integral picks out just the A_n term:

$$\int_0^L f(x)\sin\left(2\pi\frac{nx}{2L}\right)dx = 0+0+\cdots+0+A_n\frac{L}{2}+0+0+\cdots$$

Finally, this means that we can calculate the nth mode coefficient by the following equation:

$$A_n = \frac{2}{L}\int_0^L f(x)\sin\left(2\pi\frac{nx}{2L}\right)dx$$

This scheme is called "Fourier analysis." It plays an enormously important role in many branches of science and engineering.

*A Specific Example

In Section S3.5 we wanted to analyze a "square wave" as a sum of modes (Figure S3.59). This function $f(x)$ has the value $+1$ for $0 < x < L/2$, and 0 for $L/2 < x < L$, so we have

$$A_n = \frac{2}{L}\int_0^L f(x)\sin\left(2\pi\frac{nx}{2L}\right)dx = \frac{2}{L}\int_0^{L/2} f(x)\sin\left(2\pi\frac{nx}{2L}\right)dx + 0$$
$$A_n = \frac{2}{L}\left[-\frac{L}{\pi n}\cos\left(\pi\frac{nx}{L}\right)\right]_0^{L/2} = \frac{2}{\pi}\frac{1}{n}\left[1-\cos\left(n\frac{\pi}{2}\right)\right]$$

Dropping the overall factor of $2/\pi$ (which just sets the overall scale of how big the oscillation is), we get the following set of mode coefficients: $A_1 = 1$, $A_2 = 1$, $A_3 = 1/3$, $A_4 = 0$, $A_5 = 1/5$, $A_6 = 1/3$, $A_7 = 1/7$, $A_8 = 0$, $A_9 = 1/9$, $A_{10} = 1/5$, and so on.

Earlier in this section we used just these first ten modes, so we didn't mimic the desired square-wave function $f(x)$ exactly right, but we were close. If we had used more modes, we would have come even closer to approximating the square wave. If we use an arbitrarily large number of modes we can come arbitrarily close to the desired function.

This is actually a pretty extreme example, because this square wave has discontinuities that require very high-frequency modes in order to approximate the infinite slopes of the function. Functions that have less extreme behavior can often be approximated quite well with only a few modes.

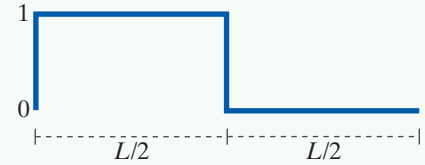


Figure S3.59 We will express this square wave in terms of a sum of sinusoids.

S3.8 *DERIVATION: TWO SLITS ARE LIKE TWO SOURCES

The proof that two illuminated slits act like two sources is an application of the superposition principle. Suppose that a laser illuminates an opaque sheet (Figure S3.60). According to the superposition principle, electric and magnetic fields produced by the laser must be present behind the sheet, despite the lack of light there. This means that the light from the laser accelerates electrons in the sheet in such a way that the accelerated electrons produce electric and

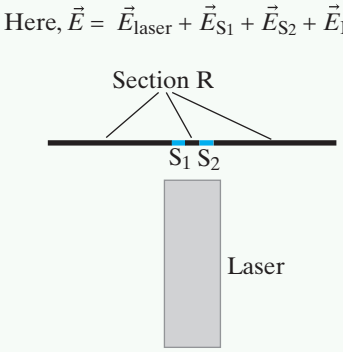


Figure S3.60 Behind the opaque sheet, the net electric field is zero.

magnetic fields behind the sheet that exactly cancel the fields produced by the laser.

We can consider the sheet to consist of the material that will eventually be removed to make the slits (sections S_1 and S_2 of the sheet), plus the rest of the sheet (section R). The (zero) net electric and magnetic fields beyond the sheet can be considered to be produced by four sources: the laser and the accelerated electrons in sections S_1, S_2, and R of the sheet. The net electric field is

$$\vec{E}_{\text{laser}} + \vec{E}_{S_1} + \vec{E}_{S_2} + \vec{E}_{R} = 0$$

When we cut out the slits, the net electric and magnetic fields beyond the sheet are due just to the laser and to the accelerated electrons in section R of the sheet, $\vec{E}_{\text{laser}} + \vec{E}_{R}$ (Figure S3.61).

If we assume that the electron accelerations in section R are about the same in the two situations, whether or not the slits have been cut out, then the electric and magnetic fields produced by section R are about the same in both situations. Since $\vec{E}_{\text{laser}} + \vec{E}_{S_1} + \vec{E}_{S_2} + \vec{E}_{R} = 0$, the electric field with the slits cut out is

$$\vec{E}_{\text{laser}} + \vec{E}_{R} = -(\vec{E}_{S_1} + \vec{E}_{S_2})$$

In calculating intensities we don't care about an overall sign, so instead of considering the difficult situation of a laser and a sheet with two slits cut out of it, we can consider the much simpler and already familiar situation of two sources, located where the slits actually are and producing a net electric field $\vec{E}_{S_1} + \vec{E}_{S_2}$, with no laser and no section R (Figure S3.62).

How good is the assumption that the electron accelerations in section R are about the same whether or not the slits have been cut out? This assumption may not work well very close to the slits, because the presence or absence of the material will affect the edge regions of section R, but we are interested in explaining the interference pattern on a distant screen, not close to the slits.

The assumption also may not work very well if the slits are too small, because in that case their contributions are so small as to be comparable to the edge effects, but we typically deal with slits that are quite wide in the sense of having a width much larger than the wavelength of light. Predictions based on treating the two slits as though they were sources do give results in excellent agreement with observations.

This derivation is due to Richard Feynman (*The Feynman Lectures on Physics*, by R. P. Feynman, R. B. Leighton, and M. Sands, Addison-Wesley 1964, page 31–10.

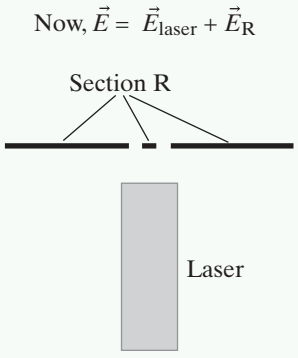


Figure S3.61 With the slits cut out, the net field is due just to the laser and section R.

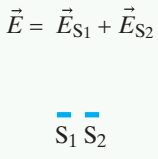


Figure S3.62 Beyond the slits, we can pretend that the field is due to just these two sources.

S3.9 *THE WAVE EQUATION FOR LIGHT

In this optional section we derive the wave equation for light, starting from the differential forms of Maxwell's equations, where $\vec{J}$ is the current density in amperes per square meter:

$$\text{div}(\vec{E}) = \vec{\nabla} \bullet \vec{E} = \frac{\rho}{\varepsilon_0} \qquad \text{Gauss's law}$$

$$\text{div}(\vec{B}) = \vec{\nabla} \bullet \vec{B} = 0 \qquad \text{Gauss's law for magnetism}$$

$$\text{curl}(\vec{E}) = \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \qquad \text{Faraday's law}$$

$$\text{curl}(\vec{B}) = \vec{\nabla} \times \vec{B} = \mu_0 \left(\vec{J} + \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \right) \qquad \text{Ampere–Maxwell law}$$

Consider the case of an electromagnetic plane wave propagating in the x direction in empty space, where there are no charges or currents, with $\vec{E} = \langle 0, E_y, 0 \rangle$ and $\vec{B} = \langle 0, 0, B_z \rangle$. Recall that

$$\vec{\nabla} = \langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \rangle$$

Using the standard definitions for calculating the components of a dot product and a cross product, you can verify that Maxwell's equations for this situation are as follows (remember that there are no charges or currents):

$$\begin{aligned}
\text{div}(\vec{E}) &= \langle 0, \frac{\partial E_y}{\partial y}, 0 \rangle = \langle 0, 0, 0 \rangle \\
\text{div}(\vec{B}) &= \langle 0, 0, \frac{\partial B_z}{\partial z} \rangle = \langle 0, 0, 0 \rangle \\
\text{curl}(\vec{E}) &= \langle -\frac{\partial E_y}{\partial z}, 0, \frac{\partial E_y}{\partial x} \rangle = \langle 0, 0, -\frac{\partial B_z}{\partial t} \rangle \\
\text{curl}(\vec{B}) &= \langle \frac{\partial B_z}{\partial y}, -\frac{\partial B_z}{\partial x}, 0 \rangle = \mu_0 \varepsilon_0 \langle 0, \frac{\partial E_y}{\partial t}, 0 \rangle
\end{aligned}$$

These equations show that E_y and B_z don't depend on y or z, so we are indeed dealing with a plane wave propagating in the $+x$ or $-x$ direction. The remaining two component equations are these:

$$\begin{aligned}
\frac{\partial E_y}{\partial x} &= -\frac{\partial B_z}{\partial t} \\
-\frac{\partial B_z}{\partial x} &= \mu_0 \varepsilon_0 \frac{\partial E_y}{\partial t}
\end{aligned}$$

Take the partial derivative of the first equation with respect to x, and reverse the order of the partial derivatives of B_z:

$$\frac{\partial}{\partial x}\left(\frac{\partial E_y}{\partial x}\right) = -\frac{\partial}{\partial x}\left(\frac{\partial B_z}{\partial t}\right) = -\frac{\partial}{\partial t}\left(\frac{\partial B_z}{\partial x}\right)$$

Substitute the value of $\partial B_z / \partial x$ from the second equation:

$$\frac{\partial^2 E_y}{\partial x^2} = \mu_0 \varepsilon_0 \frac{\partial^2 E_y}{\partial t^2}$$

Rewrite to match the format of the wave equation:

$$\frac{\partial^2 E_y}{\partial t^2} = \frac{1}{\mu_0 \varepsilon_0} \frac{\partial^2 E_y}{\partial x^2}$$

This is the equation of a wave moving with speed $v = 1/\sqrt{\mu_0 \varepsilon_0}$, which evaluates to $v = 3 \times 10^8$ m/s. Recall that we obtained the same result in Chapter 23 by applying the integral forms of Maxwell's equations to a slab of electromagnetic radiation.

SUMMARY

Light has both wave-like and particle-like properties. A photon is a particle of light with zero rest mass, with energy and momentum related to its wavelength:

$$\lambda = \frac{h}{p} \quad \text{and} \quad E = \frac{hc}{\lambda}$$

Both photons and particles with nonzero rest mass such as electrons and neutrons can exhibit wavelike properties, such as interference.

Properties of sinusoidal waves

A wave propagating to the right: $E\cos(2\pi\frac{t}{T} - 2\pi\frac{x}{\lambda} + \phi)$

E is the amplitude.
T is the period.
$f = 1/T$ is the frequency.
$\omega = 2\pi f$ is the angular frequency.
λ is the wavelength.
ϕ is a phase corresponding to choice of $t = 0$ and $x = 0$.
$v = f\lambda$ (v is the speed of propagation of the wave).
Intensity I is proportional to (amplitude)^2.

Interference

Waves of the same frequency can interfere with each other. As a result of the superposition principle, some locations have unusually large amplitudes and intensities, whereas other locations have small or zero amplitudes and intensities.

Two-slit interference may be observed even if photons go through the slits one at a time.

Standing waves

Standing waves result from the interference of two traveling waves of the same frequency that are traveling in opposite directions. The standing waves have the same frequency and wavelength as the traveling waves.

In a confined space standing waves are quantized—only certain standing-wave wavelengths are allowed. In the simple case of a taut string, the allowed wavelengths are $2L$, $2L/2$, $2L/3$, and so on.

The standing-wave motions are called "modes." Arbitrarily complex confined waves can be built up in terms of superpositions of these simple modes.

The steady state is reached when the energy losses are just equal to the input energy.

Path difference and two-source interference

Interference of two in-phase sources that are a distance d apart:

Maximum intensity ($4I_0$) where $\Delta l = d\sin\theta$ is 0, λ, 2λ, 3λ, and so on.

Minimum intensity (zero) where $\Delta l = d\sin\theta$ is $\lambda/2$, $3\lambda/2$, $5\lambda/2$, and so on.

No zero-intensity minimum is possible if $d < \lambda/2$.

If the sources are not in phase, you need to take that phase difference into consideration in addition to the phase difference due to Δl.

X-ray diffraction

Condition for "reflection" is $2d\sin\theta = n\lambda$, where n is an integer, d is the distance between adjacent crystal planes, and θ is the angle of the incoming x-rays to the surface. Polycrystalline powders produce rings around the beam.

Compton scattering prediction

$$\lambda_2 - \lambda_1 = \frac{h}{mc}(1 - \cos\theta)$$

Thin films show interference effects. Examples are oil slicks and soap bubbles.

Diffraction gratings

Maxima at angles satisfying $d\sin\theta = n\lambda$, where d is the slit spacing.

Angular resolution of devices

$$\Delta\theta \approx \frac{\lambda}{W}$$

where W is the total width of the (illuminated portion) of the device.

Single source of width W: angle to first diffraction minimum is $\theta \approx \lambda/W$.

Scattering of light is affected by interference effects.

The wave equation: $\dfrac{\partial^2 u(x,t)}{\partial t^2} = v^2\dfrac{\partial^2 u(x,t)}{\partial x^2}$

The speed of longitudinal waves: $v = \sqrt{\dfrac{k_{s,i}}{m_a}}\,d = \sqrt{\dfrac{Y}{\rho}}$

The speed of transverse waves: $v = \sqrt{\dfrac{F_T d}{m_a}} = \sqrt{\dfrac{F_T}{\mu}}$

From allowed wavelengths in a standing wave one can calculate the corresponding frequencies. In the simple case of a taut string, the allowed frequencies can be calculated from $f = v/\lambda$, and they are $v/(2L)$, $2v/(2L)$, $3v/(2L)$, and so on.

Particles have wave-like properties

The momentum of a particle is related to its wavelength:

$$\lambda = \frac{h}{p}$$

where $h = 6.6 \times 10^{-34}$ J-s is Planck's constant.

The photoelectric effect

A photon with energy equal to or greater than the binding energy of an electron and a metal (the "work function") can eject an electron from a metal surface.

Compton scattering

A collision between a photon and a free electron results in a change of the wavelength of the photon, corresponding to the amount of energy gained by the electron in the collision.

Fourier analysis

A confined wave can be expressed as a sum of sinusoids:

$$f(x) = A_1 \sin\left(2\pi\frac{x}{2L}\right) + A_2 \sin\left(2\pi\frac{2x}{2L}\right) + A_3 \sin\left(2\pi\frac{3x}{2L}\right) + \cdots$$

where

$$A_n = \frac{2}{L}\int_0^L f(x)\sin\left(2\pi\frac{nx}{2L}\right)dx$$

QUESTIONS

Q1 It is natural to say that "light bounces off mirrors," but is there a physics principle that would account for such behavior? What really happens? Why isn't there light going in all directions, not just in the direction of the "reflection" angle?

Q2 Summarize the different predictions of the wave and particle models of light regarding the photoelectric effect. What experimental observations support the particle model?

Q3 According to the wave model of light, what is the relationship between energy and intensity? What is it according to the particle model of light?

Q4 Figure S3.63 is an intensity plot for a situation where the sources are close together, with $d = \lambda/3.5$. We see that the intensity distribution isn't very dramatic when $d < \lambda/2$. Why is the intensity nonzero everywhere? Why isn't the intensity zero somewhere?

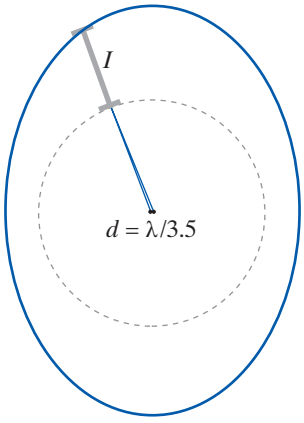


Figure S3.63 Intensity plot for two sources very close together.

Q5 In a collision between a photon and a stationary electron, why does the wavelength of the photon change as a result of the collision? Does it increase or decrease?

Q6 In electron diffraction, diffraction rings are produced when electrons go through polycrystalline material after being accelerated through an accelerating potential difference. What happens to the size of the diffraction rings when the accelerating potential difference is increased?

Q7 Explain why a beam of light can go straight through a rather long tank of clear water or a long rod of clear glass, with hardly any light emitted to the side despite the huge number of atoms whose electrons are accelerated by the incoming light.

Q8 At night you look down into an outdoor swimming pool that has a powerful underwater light whose horizontal beam heads to your right. The water is not completely clear, and there is significant scattered light. You have a polarizing film with a line drawn on it showing the direction of electric field that is passed through with little loss. To minimize the brightness of the scattered light, should you hold the film with the line in the direction of the beam (left/right) or perpendicular to the beam (up/down)? Explain briefly, including a diagram.

Q9 A satellite in Earth orbit carries a camera. Explain why the diameter of the camera's lens determines how small an object can be resolved in a photo taken by the camera.

Q10 How do standing waves differ from traveling waves? How are they similar?

Q11 When a cello string is bowed, many different modes are excited, not just the lowest-frequency mode, which contributes to the richness of the sound. **(a)** When the bow is taken away the string continues to vibrate. After some time, only the lowest-frequency mode remains in motion. Why? **(b)** When a cellist bows while lightly touching the midpoint of the string, without pressing the finger down on the string, the lowest-frequency mode and all modes with odd multiples of that frequency are absent. This changes the tone markedly and is used to achieve a special musical effect. Explain why the odd harmonics are missing.

PROBLEMS

Section S3.1

•P12 A particular AM radio station broadcasts at a frequency of 1020 kHz. What is the wavelength of this electromagnetic radiation? How much time is required for the radiation to propagate from the broadcasting antenna to a radio 4 km away?

•P13 The wavelength of violet light is about 400 nm (1 nm = 1×10^9 m). What are the frequency and period of the light waves?

•P14 The height of a particular water wave is described by the function

$$y = (1.5\text{ m})\cos\left[(25.1/\text{s})t - (2.51/\text{m})x + (\frac{\pi}{2}\text{ rad})\right]$$

Calculate the angular frequency ω, the frequency f, the period T, the wavelength λ, the speed and direction of propagation (the $+x$

or the $-x$ direction), and the amplitude of the wave. At $x = 0$ and $t = 0$, what is the height of the wave?

•P15 The relations between wavelength, frequency, period, and speed of propagation apply to all wave phenomena, not just electromagnetic radiation. Middle C on a piano has a fundamental frequency of about 256 Hz. What is the corresponding wavelength of the sound waves? (The speed of sound in air varies with temperature but is about 340 m/s at room temperature.)

•P16 Two audio speakers are side by side, 1 m apart. They are connected to the same amplifier, which is producing a sine wave of 440 Hz ("concert A"). Calculate a direction in which you won't hear anything, and make a diagram showing the speakers and this direction. (The speed of sound in air is about 340 m/s.)

••P17 In Figure S3.64 an unpolarized sinusoidal electromagnetic wave with wavelength λ travels along the $-x$ direction in a region where there are two short copper wires oriented along the z direction, a distance $L = 2.5\lambda$ apart.

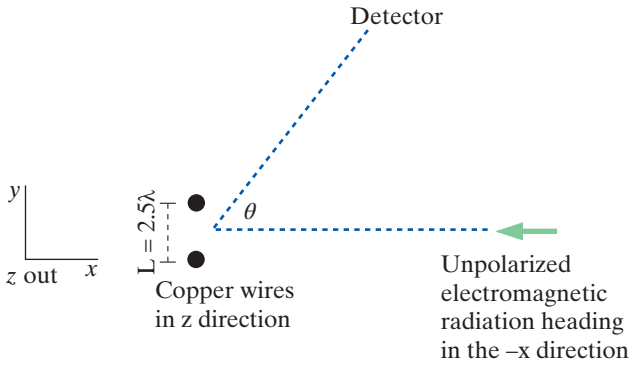


Figure S3.64

(a) You place a detector a long ways away, at an angle θ to the x axis. Despite being far outside the region of the unpolarized wave, you detect electromagnetic radiation, and it is polarized. Explain this phenomenon in detail, including the directions of the electric and magnetic fields that you detect. **(b)** Calculate an angle θ_1 other than 0° where you will see maximum intensity, and calculate an angle θ_2 where you will see zero intensity. Explain briefly.

••P18 Coherent green light with a wavelength of 500 nm illuminates two narrow vertical slits a distance 0.12 mm apart. Bright green stripes are seen on a screen 2 m away. **(a)** How far apart are these stripes, center-to-center? **(b)** Do they get farther apart or closer together if you move the slits closer together? **(c)** Do they get farther apart or closer together if you use violet light (wavelength = 400 nm)?

Section S3.2

•P19 A ray of violet light (wavelength 400 nm) hits perpendicular to a transmission diffraction grating that has 10,000 lines per centimeter. At what angles to the perpendicular are there bright violet rays, in addition to zero degrees?

•P20 At night you look through a transmission diffraction grating at a sodium-vapor lamp used for outdoor lighting, which emits nearly monochromatic light of wavelength 588 nm. The manufacturer of the grating states that it was ruled with 10,000 lines per centimeter. In addition to zero degrees, at what other angles will you observe bright light?

•P21 When x-rays with wavelength 0.4×10^{-10} m hit a crystal at an angle of 30° to the surface, a strong beam of x-rays is observed in the direction of the "reflection" angle. Assume that you know from other evidence that the spacing of the atomic layers parallel to this surface of the crystal is greater than 1×10^{-10} m and less than 1.8×10^{-10} m. What are possible values of the spacing between the layers?

••P22 Suppose that you have an x-ray source that produces a continuous spectrum of wavelengths, with the shortest wavelength being 0.3×10^{-10} m. You have a crystal whose atoms are known to be arranged in a cubic array with the distance between nearest neighbors equal to 1.2×10^{-10} m. **(a)** Design an arrangement of x-ray beam and crystal orientation that will give you a monochromatic beam of x-rays whose wavelength is 0.5×10^{-10} m. Explain your arrangement carefully and fully in a diagram. **(b)** If the shortest wavelength in the x-ray spectrum were 0.2×10^{-10} m instead of 0.3×10^{-10} m, explain why your beam would not be a pure single-wavelength beam.

••P23 The color magenta consists of a mixture of red (about 700 nm) and blue light (about 450 nm). If you see a magenta-colored section of a soap bubble, about how thick is this section of the soap bubble? Assume that wavelengths in the material are shortened by a factor of 1.3.

Section S3.3

•P24 Red laser light with wavelength 630 nm goes through a single slit whose width is 0.05 mm. What is the width of the image of the slit seen on a screen 5 m from the slit?

••P25 A single vertical slit 0.01 mm wide is illuminated by red light of wavelength 700 nm. **(a)** About how wide is the bright stripe on a screen 2 m away? **(b)** Does the stripe get wider or narrower if you make the slit narrower? **(c)** Does the stripe get wider or narrower if you use blue light?

••P26 Consider a communications satellite that orbits the Earth at a height of about 40,000 km. At this height the orbit period is 24 h, so that it seems to hang motionless above the Earth, which turns underneath the satellite once every 24 h. This is called a "synchronous" satellite. The advantage of such an arrangement is that receiving antennas on the ground can point at a seemingly fixed location of the satellite in the sky and not have to be steered. You have probably seen such fixed receiving antennas pointed at the sky. If the broadcasting antenna on the satellite is 2 m in diameter and broadcasts at a frequency of 10 GHz ($f = 10 \times 10^9$ Hz), approximately what is the diameter of the region on the ground where the satellite transmission can be picked up?

Section S3.4

•P27 A rope 4 m long with a mass of 0.06 kg is stretched taut. A video shows that a transverse pulse takes 0.16 s to run the length of the rope. What is the tension in the rope?

•P28 Strike the end of a metal bar that is 2 m long, simultaneously starting an electronic timer, which is then stopped when a microphone at the other end of the bar first detects a sound. The measured time for sound to go from one end of the bar to the other is 0.6 ms, and the density of the bar is measured to be 7.4 g/cm^3. What is Young's modulus for this material?

•P29 In Figure S3.65, what is the amplitude? What is the wavelength? What is the period? What is the wave number? What is the angular frequency? What is the frequency?

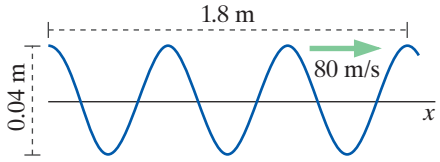


Figure S3.65

•P30 Write the function for a cosine wave moving in the $-x$ direction along a rope with speed 30 m/s and amplitude 0.04 m.

•P31 Write the function in terms of x and t for a cosine wave moving in the $-x$ direction along a rope with wavelength 0.2 m, speed 45 m/s, and amplitude 0.04 m.

•P32 A sinusoidal wave moves along a vertical rope, with $u(y,t) = 0.03\sin(20y + 45t)$. What is the speed of propagation of the wave? What is the frequency? Is the wave moving upward or downward?

••P33 One end of a metal bar is driven sinusoidally in a longitudinal mode with frequency 800 Hz. At the other end of the bar the sound intensity is measured to be 1×10^{-12} W/m^2, which is approximately the smallest sound intensity the human ear can detect. The bar is 4 cm long, with cross-sectional area 1.3 cm by 0.7 cm. The density is 7 g/cm^3 and the speed of sound is 1700 m/s. What is the amplitude of the longitudinal oscillations of an atom in the bar? How does this compare with typical interatomic distances in a solid?

••P34 A wire of total mass M and length L hangs from the ceiling, under its own weight. How long does it take for a transverse pulse to travel from the bottom of the wire to the top? Be sure to show all of the steps in your analysis.

Section S3.5

•P35 The lowest-frequency mode of the lowest (C) string on a cello is 64 Hz. The length of the string between its supports is 70 cm. **(a)** What is the speed of propagation of traveling waves on this string? **(b)** Calculate the frequencies and wavelengths of the next four modes (not including the 64 Hz), and make a sketch of the standing-wave shape for each mode.

Section S3.6

•P36 What is the energy in eV of a photon whose wavelength is 334 nm?

•P37 If the work function of a metal is 3.4 eV, what would be the maximum wavelength of light required to eject an electron from the metal?

•P38 Electrons are accelerated through a potential difference of 1000 V and strike a polycrystalline powder whose atomic layers are 1×10^{-10} m apart. Predict an angle at which you will see an electron-diffraction ring.

••P39 A 100 W light bulb is placed in a fixture with a reflector that makes a spot of radius 20 cm. Calculate approximately the amplitude of the radiative electric field in the spot and the number of photons per second hitting the spot.

ANSWERS TO CHECKPOINTS

1 When $t = 0$, we have $E\cos(\pi) = -E$, which is one point on the curve. The curve is shifted π rad (a half cycle) relative to the first curve (Figure S3.66).

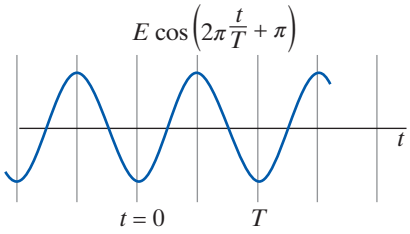


Figure S3.66

2 9
3 **(a)** 6.4° (0.11 rad), **(b)** 63° (1.1 rad)
4 13.89°
5 640 nm
6 588 nm
7 8 mm
8 1.2 m
9 We assumed that the slope of $u(x)$ was small in order to satisfy the small-angle approximation for sine and cosine.
10 The long-wavelength (low-frequency) mode should be easier, because lower speeds mean lower air resistance, and longer wavelength means smaller bending angles (lower internal friction).
11 3.4 eV
12 5.24×10^{-11} m
13 About 150 V for $\lambda \approx 1 \times 10^{-10}$ m